

# A comprehensive model of the quiet-time, near-Earth magnetic field: phase 3

Terence J. Sabaka,<sup>1</sup> Nils Olsen<sup>2</sup> and Robert A. Langel<sup>3\*</sup>

<sup>1</sup>Raytheon ITSS, Lanham, MD 20706, USA. E-mail: sabaka@geomag.gsfc.nasa.gov

<sup>2</sup>Danish Space Research Institute, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark

<sup>3</sup>Geodynamics Branch, Code 921, NASA/GSFC, Greenbelt, MD 20771, USA

Accepted 2002 February 15. Received 2001 September 28; in original form 2000 September 6

## SUMMARY

The near-Earth magnetic field is caused by sources in the Earth's core, ionosphere, magnetosphere, lithosphere and from coupling currents between the ionosphere and the magnetosphere, and between hemispheres. Traditionally, the main field (low degree internal field) and magnetospheric field have been modelled simultaneously, with fields from other sources being modelled separately. Such a scheme, however, can introduce spurious features, especially when the spatial and temporal scales of the fields overlap. A new model, designated CM3 (Comprehensive Model: phase 3), is the third in a series of efforts to coestimate fields from all of these sources. This model has been derived from quiet-time Magsat and POGO satellite and observatory hourly means measurements for the period 1960–1985. It represents a significant advance in the treatment of the aforementioned field sources over previous attempts, and includes an accounting for main field influences on the magnetosphere, main field and solar activity influences on the ionosphere, seasonal influences on the coupling currents, *a priori* characterization of the influence of the ionosphere and the magnetosphere on Earth-induced fields, and an explicit parametrization and estimation of the lithospheric field. The result is a model that describes well the 591 432 data with 16 594 parameters, implying a data-to-parameter ratio of 36, which is larger than several popular field models.

**Key words:** Earth's magnetic field, electromagnetic induction, geomagnetic variation, ionosphere, lithosphere, magnetosphere.

## 1 INTRODUCTION

The terrestrial magnetic field is composed of contributions from many sources. Resolution of these constituent fields is of primary importance in understanding the physical processes responsible for their existence. This paper is concerned with the modelling of the field from measurements at the Earth's surface and extending to about 2000 km in altitude above that surface, a region that will be referred to as 'near-Earth'. Before embarking on this inverse problem, however, a brief background into the nature of these source phenomena is in order.

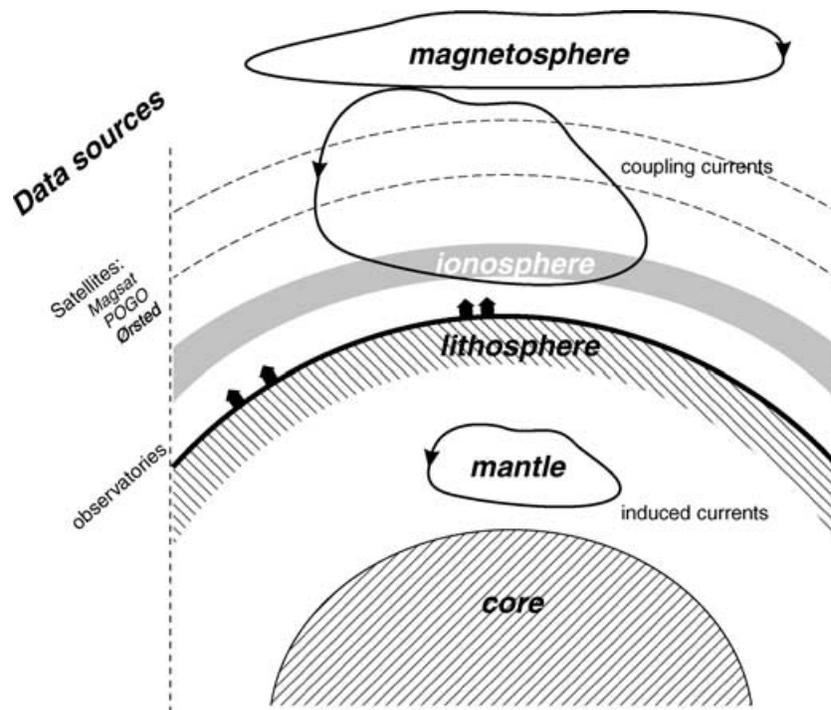
### 1.1 Near-Earth magnetic field contributions

A schematic of the current sources contributing to the near-Earth magnetic field is given in Fig. 1. By far the most dominant of these fields is of core origin, accounting for over 97 per cent of the field observed at the Earth's surface and ranging in intensity from about 30 000 nT at the equator to about 50 000 nT at the poles. Accord-

ing to geodynamo theory, inductive interactions between the fluid motion of the liquid outer core and the geomagnetic field not only modify the source current so as to induce secular variation (SV) of the field, but sustain it against long-term decay caused by magnetic diffusion and Ohmic dissipation of the source current (see Hollerbach 1996). The timescale of this field is of the order of centuries with typical SV magnitudes at the core–mantle boundary (CMB) of a few thousand nT yr<sup>-1</sup> (see Jackson *et al.* 2000).

The solar-quiet (Sq) magnetic field variation is a manifestation of an ionospheric current system. Heating on the dayside and cooling on the nightside of the atmosphere generates tidal winds that drive ionospheric plasma against the geomagnetic field, inducing electric fields and currents in the dynamo region between 100–140 km in height. The current system remains relatively fixed to the Earth–Sun line and produces regular, broad-scale daily variations that are seen directly in the magnetograms of geomagnetic 'quiet' days, hence the name Sq. On 'disturbed' days there is an additional variation, which includes superimposed magnetic storm signatures. Because the geomagnetic field is strictly horizontal at the dip equator, there is an enhancement of the effective conductivity, which results in an enhanced eastward current, called the equatorial electrojet (EEJ), flowing along a few hundred kilometre wide band centred on the

\*Deceased, 2000 February 9.



**Figure 1.** Schematic of current sources contributing to the near-Earth magnetic field and data sources considered in this study.

dayside dip equator (see Onwumechilli & Ozoemena 1989). In addition, auroral electrojets (AEJ) flow in the auroral belt and vary in amplitude with different levels of magnetic activity. At the Earth's surface, the Sq fields are of the order of 10–50 nT, depending upon component, latitude, season, solar activity and time of day; the magnetic signature of the EEJ can be about 5–10 times that of Sq; and that of the AEJ can vary widely from a few 10s nT during quiet periods to several thousand nT during major magnetic storms (see Volland 1984).

The field originating in the Earth's magnetosphere is a result primarily of the ring-current and the currents on the magnetopause and in the magnetotail. Currents flowing on the outer boundary of the magnetospheric cavity, known as the magnetopause, cancel the Earth's field outside and distort the field within the cavity. This produces an elongated tail in the antisolar direction within which sheet currents are established in the equatorial plane. The interaction of these currents with the radiation belts near the Earth produces a ring-current in the dipole equatorial plane, which partially encircles the Earth, but achieves closure via field-aligned currents (FACs) into and out of the ionosphere. These resulting broad-scale fields have magnitudes of the order of 20–30 nT near the Earth during magnetically quiet periods, but can increase to several hundred nT during disturbed times (see Olson 1984).

If displacement currents are neglected, then the current densities associated with these external fields are solenoidal and therefore must flow along closed circuits. Given the complex nature of the conductivity structure in the near-Earth region, circuit closure is sometimes achieved through currents that couple the various source regions. At high latitudes, the auroral ionosphere and magnetosphere are coupled by currents that flow along the Earth's magnetic field lines (see Potemra 1982). The fields from these FACs have magnitudes that vary with the magnetic disturbance level. However, they are always present (of the order of 30–100 nT during quiet periods and up to several thousand nT during substorms). Fields from these

currents have been detected in surface data in the *Y* (east) component of the magnetic field at low latitudes, with difficulty, but are mostly mapped using magnetometers aboard near-Earth-orbiting satellites. There are also currents that couple the Sq currents systems in the two hemispheres that flow, at least in part, along magnetic field lines. Detection of these has been reported by Olsen (1997a) using data from the Magsat satellite. The associated magnetic fields are generally 10 nT or less. Finally, there exists a meridional current system that is connected to the EEJ with upward-directed currents at the dip equator and field-aligned downward-directed currents at low latitudes (within 15° of the dip equator). Fields from this current system have been detected by magnetic measurements taken on a rocket (Musmann & Seiler 1978) and from those taken by Magsat (Maeda *et al.* 1982). In the latter case, the EEJ coupling currents resulted in fields of about 15–40 nT in the Magsat data at dusk local time.

The lithosphere is a rheological classification for that outer layer of the Earth, which is rigid and the crust is the petrologically distinct upper portion of this. The lithosphere contains regions for which the temperature is below the Curie point of magnetite and other magnetic minerals. As a result, it can have magnetization that is either induced by the present-day ambient field or frozen into the rocks at their last time of cooling below the Curie temperature, i.e. remanent magnetization. Fields from the lithosphere are of amplitude up to several thousand nT at the surface and at aircraft altitude, and up to about 20 nT at the satellite altitudes considered in this paper (see Langel & Hinze 1998). The length-scales of lithospheric fields range from extremely broad, e.g. continent–ocean magnetization contrast (Counil *et al.* 1991), to very local, although only features larger than a few hundred kilometres may be resolved at satellite altitude.

The time-varying external fields induce currents in the Earth's electrically conducting interior. These currents produce a secondary magnetic field, the size of which depends upon the amplitude of the inducing field, the mantle conductivity and the period. For realistic

values of mantle conductivity and induction on a global scale, the amplitude of the induced contribution is about 40 per cent of the inducing amplitude for periods of a few hours, about 30 per cent for periods of a few days and less than 20 per cent for periods of more than 6 months (see Parkinson & Hutton 1989).

### 1.2 Earlier modelling efforts

Historically, fields from the various sources have been modelled separately, or at least not all together. Under the assumption that measurements are acquired in current-free regions, models of the core, magnetospheric, ionospheric and lithospheric fields have taken the form of gradients of Laplacian potential functions, usually in spherical coordinates. In their quest to map the magnetic field through time, Bloxham & Jackson (1992) and later Jackson *et al.* (2000), with an extension to earlier epochs, omit fields external to satellite sampling regions and include the effects of crustal and ionospheric fields as noise sources while constructing smooth spherical harmonic models of the core SV at the CMB from satellite, observatory first-difference and survey data. However, these excluded external fields, presumably of magnetospheric origin, may be coestimated with internal fields. This approach has been used by Langel & Estes (1985a,b) to map the main field (the ‘low’-degree spherical harmonic contributions from both the core and the lithosphere) in 1980 from Magsat satellite data corrected for ionospheric effects. They include an external field with an associated induced contribution of spherical harmonic  $Y_1^0$ , the time variations of which are proportional to the  $D_{st}$  index. Sabaka *et al.* (1997) followed a similar method in modelling long-period variation of the main and magnetospheric fields using measurements from satellites, observatory annual-means (OAMs) and survey data, but with annual averages of the *aa* index (Mayaud 1972, 1980; Rangarajan 1989) as a proxy for the  $D_{st}$  index. This coestimation of main, magnetospheric and induced fields has also been applied to nightside Ørsted satellite data, culminating in the Ørsted Initial Field Model (OIFM) of Olsen *et al.* (2000).

Spherical harmonic models of ionospheric fields have generally been produced separately from the other field contributions using data from magnetic observatories and variometer stations (see e.g. Matsushita & Maeda 1965; Malin 1973; Winch 1981; Campbell 1989; Olsen 1997b; Schmucker 1999). Langel & Estes (1985a) reported detection of Sq fields in the data from the POGO (Polar Orbiting Geophysical Observatories, comprising the OGO-2, OGO-4 and OGO-6 satellites) satellites. Attempts to model the EEJ effects in satellite data directly have been carried out by Langel *et al.* (1993), who first isolate the dip-latitude-dependent fields via filtering and then fit with either zonal harmonics in dipole coordinates or other related empirical functions.

Global models of the lithospheric field are realizable only with satellite data, and have taken the form of various potential field representations. The usual approach has been to isolate the lithospheric field first by removing estimates of the main, ionospheric and magnetospheric fields from the data and then to correlate what is believed to be the remaining signal (see Langel & Hinze 1998, for details on recommended procedures). Though the exact methods may deviate from this, some examples of studies of this type can be found in Arkani-Hamed & Strangway (1985a,b, 1986), Arkani-Hamed *et al.* (1994), Cohen & Achache (1990), Counil *et al.* (1991), Hamoudi *et al.* (1998), Ravat *et al.* (1995). A natural alternative is to include higher-degree/higher-order terms in the internal field potential expansion to capture the lithospheric contributions. Cain *et al.* (1984) followed this procedure in deriving a degree/order 29 internal field model from Magsat data corrected for both magnetospheric and

ionospheric effects, and later, Schmitz *et al.* (1989) and Cain *et al.* (1989a, 1990) derive even higher-degree ( $\geq 50$ ) expansions using improved ionospheric data corrections.

Currents that couple the ionosphere with the magnetosphere flow in the ionospheric *F*-region at satellite altitude and therefore produce toroidal magnetic fields (see Backus 1986). Takeda & Maeda (1983) modelled the fields caused by meridional currents as an *F*-region dynamo, but perhaps the best global model of the fields from these coupling currents comes from the work of Olsen (1997a), who represents them as a toroidal stream function expansion within the Magsat sampling shell under the assumption of radial currents only.

### 1.3 Comprehensive approach

Each of the studies cited in the previous section enjoy varying degrees of success in their ability to describe the target source field. They all, however, suffer at some level from the effects of frequency overlap between the spectra of the various source fields, both in the spatial and temporal domains. That is, frequency range cannot be used to absolutely distinguish between the spectra, and so bandpass and bandstop filters alone are doomed to either keep some of the unwanted signal or eliminate some of the signal of interest. Note that this is different from aliasing, which is imposed by sampling intervals and results in signals with frequencies above the Nyquist frequency being overlapped on to those at and below it (Kanasewich 1981).

Evidently, additional information is needed to resolve the source contributions to near-Earth fields in a realistic manner. One possibility is to consider the radial positions of the various source regions relative to the available data, which is shown in the schematic in Fig. 1. Core, lithospheric and induced fields would be internal to both surface and satellite data while the magnetospheric field would be external. The ionospheric field would be external to surface data, but internal to satellite data. Thus, surface data could separate ionospheric and magnetospheric from core, lithospheric and induced fields and satellite data could separate magnetospheric from core, lithospheric, induced and ionospheric fields. This suggests that a joint analysis of both surface and satellite data could theoretically resolve parametrizations between magnetospheric, ionospheric and internal sources, but only if the parameter set is treated consistently between data types, which implies that they be coestimated. Of course, additional information would be needed to separate the various internal fields, e.g. space and timescale differences, *a priori* mantle conductivity models, etc. This simultaneous inversion for parameters describing all sources will be termed the ‘comprehensive approach’, and models of this kind could provide the reference fields needed in more refined studies where the separation/removal of the various source contributions is an issue.

This paper reports on the third in a series of efforts to derive progressively more sophisticated models using the comprehensive approach, and many of the details omitted here are included in the associated NASA Technical Memorandum of Sabaka *et al.* (2000). The first phase, reported by Sabaka & Baldwin (1993), culminated in a model known as GSFC(12/93), while the second phase, reported by Langel *et al.* (1996), culminated in the GSFC(8/95-SqM) model. These models are based upon magnetically quiet data from the POGO and Magsat spacecraft and OAMs and observatory hourly means (OHMs). Because of their limited scope, the Magsat vector data poleward of  $\pm 50^\circ$  dipole latitude are not used in order to avoid the auroral FACs, while the Magsat dusk data are corrected for the effects of the EEJ and associated meridional currents. Details of the

model parametrizations are not given here, but qualitatively, both include representations of the main field, its SV and the local time (Sun-synchronous) modes of the magnetospheric and ionospheric fields; both include ring-current variations through the  $D_{st}$  index; and both include explicit parametrizations for fields induced by the time-varying external fields. The GSFC(8/95-SqM) also includes seasonal variations in the magnetospheric and ionospheric fields.

The scope of the model presented in this paper is much wider and its attention to detail much higher than its predecessors. Being the culmination of the third phase of work along these lines, this model will be designated as ‘CM3’ (denoting Comprehensive Model: phase 3). Its aim is to describe finer systematic detail present in the data sets with a physically plausible model of least complexity, which implies a more proper distribution of the signal amongst previously considered sources and amongst newly parametrized sources, i.e. coupling currents and small-scale lithosphere. In addition to these new sources, other improvements have been made such as a smooth main field SV; a high-resolution, high-efficiency ionospheric parametrization with physical conductivity constraints and main field interaction; a magnetospheric parametrization with main field interaction; and an implicit parametrization of the induced fields associated with the time-varying external fields via a radially varying *a priori* mantle conductivity model. In the remainder of the paper, a description of the data, parametrization and method of estimation leading to the CM3 model will be given followed by a discussion of the model properties in terms of consistency and physical plausibility. In conclusion, model availability will be discussed and future directions will be outlined.

#### 1.4 Coordinate systems

It is convenient at this point to introduce the various spatial and temporal coordinate systems used throughout this paper. There are basically three spatial systems employed for different parts of the model. The usual geographic spherical polar coordinates,  $(r, \theta, \phi)$ , are defined with the  $z$ -axis through the north geographic pole,  $x$  through Greenwich and  $y$  completing the right-handed system, and are used for internal field representations. The dipole spherical polar coordinates,  $(r, \theta_d, \phi_d)$ , are defined with respect to a tilted dipole-moment vector, which in this study is supplied by the GSFC(12/83) model of Langel & Estes (1985b). The  $z$ -axis points along the moment vector out of the northern geographic hemisphere, the  $x$ -axis is perpendicular to  $z$  in the half-plane containing the south geographic pole and  $y$  completes the right-handed system. This system is used primarily to represent distant external (i.e. magnetospheric) field contributions. Finally, the ‘quasi-dipole’ (QD) coordinates proposed by Richmond (1995),  $(h_q, \theta_q, \phi_q)$ , are also tied to a magnetic field model, which in this study is the DGRF 1980 model (IAGA Division I Working Group 1 1981). The QD longitude,  $\phi_q$ , is the dipole longitude of the apex (highest point above the Earth’s surface, given by  $h_A$ ) of the magnetic field line passing through the point. The QD latitude,  $\lambda_q$ , is given by

$$\lambda_q = \pm \arccos \left( \frac{R_E + h_q}{R_E + h_A} \right)^{1/2}, \quad (1)$$

where  $R_E = 6371.2$  km is the mean radius of the Earth and  $h_q$  is altitude. Thus,  $\lambda_q$  is defined as the latitude of the footprint of an axial dipole field line through the apex, and goes to zero at the magnetic equator at all altitudes. Hence QD coordinates are defined by tracing the field line upward using DGRF 1980, and downward again using an axial dipole. This coordinate system is useful in describ-

ing ionospheric and coupling currents that are strongly horizontally organized by the ambient magnetic field.

As for the temporal coordinates, there are basically two renderings of time used. The first is the familiar universal time (UT), which tracks secular time at Greenwich and is used to describe long-term and seasonal variations in the fields. The other is magnetic local time (MLT),  $t_{mlt}$ , which is defined for an observer as

$$t_{mlt} = (180^\circ + \phi_{d,o} - \phi_{d,s})/15, \quad (2)$$

where if the dipole longitude of the observer,  $\phi_{d,o}$ , and the sub-solar point,  $\phi_{d,s}$ , are in degrees, then  $t_{mlt}$  is in hours. Again, the tilted dipole from the GSFC(12/83) model is used to assign  $t_{mlt}$  in this study. Note that the time angle is now about the dipole  $z$ -axis instead of the geographic  $z$ -axis, as in UT, and that a point fixed in MLT is Sun-synchronous. This is useful for describing external fields whose current systems exhibit diurnal cycles and are also influenced by Earth-fixed internal fields. For modelling purposes, diurnal variability is represented by functions expanded in magnetic universal time (MUT),  $t_{mut}$ , which is simply the MLT at the dipole prime meridian, i.e. set  $\phi_{d,o} = 0$  in eq. (2).

## 2 DATA

The accuracy of models derived from inverse problems is intimately related to the quality of the data being analysed. Undoubtedly some of the best data to date for the purposes of this study come from the Magsat and POGO satellite missions and from the permanent magnetic observatories, which will now be discussed.

### 2.1 Observatories

The CM3 model incorporates two samplings of OHMs from permanent magnetic observatories. The first, designated OHM-1AM, are the OHM values for which the MLTs are closest to 1 am, within 4 h, chosen from the magnetically quietest day of each month, as defined by the  $K_p$  level, within the interval 1960–1985. These data were furnished by the World Data Center for Geomagnetism in Copenhagen, Denmark, and were acquired at a somewhat latter date than the second sampling type in order to replace the OAMs first used in the model. These offer control of the main field SV between the Magsat and POGO mission duration envelopes, which is discussed further in Section 3.1.

The second, designated OHM-MUL, are chosen from the magnetically quietest day of each month during the operational periods of the POGO (1965 September to 1971 August) and Magsat (1979 November to 1980 April) missions, though data through 1982 are also included. These data are furnished by the National Geophysical Data Center (NGDC) in Boulder, CO, with augmented data from Winch, Faynberg and Singer, and others (priv. comm.). Because the shortest timescale considered in the ionospheric and magnetospheric portions of the CM3 model is 6 h (see Sections 3.2 and 3.3), only OHMs from every other hour are used. Before the OHM-MUL data set was actually analysed in the CM3 model, it underwent an outlier rejection phase by visual inspection with respect to a model derived in the preliminary stages of this study. This process is discussed in more detail in Section 4.2.

Station breaks, introduced by Langel *et al.* (1982), are times at which jumps in the observatory baselines can occur. They are assigned by a visual inspection of the time-series of vector field components. These breaks will usually coincide with a physical change in the nature of the measurements, such as a change in location,

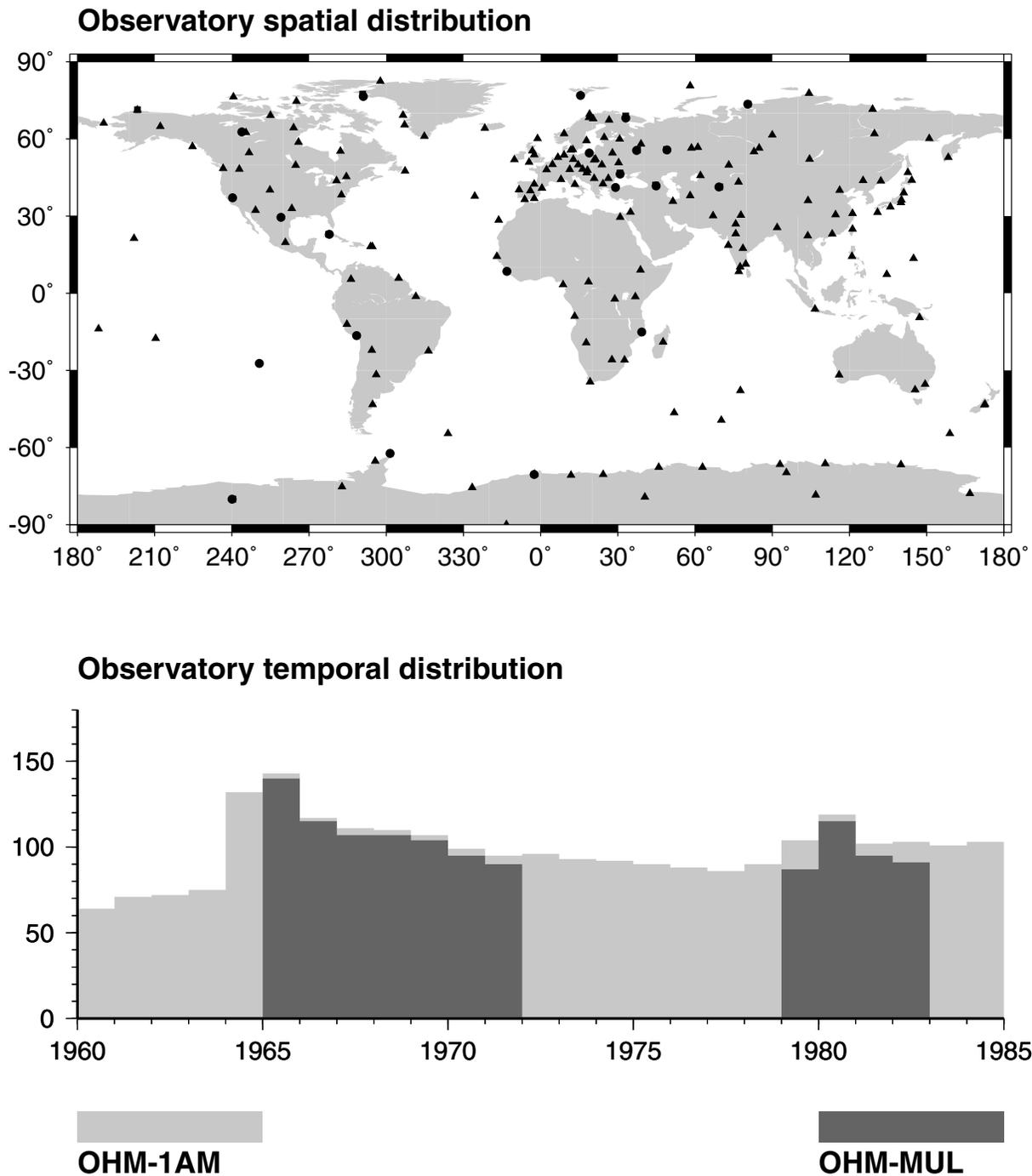
equipment or local man-made fields. Thus, each station segment may be thought to have its own baseline, which is estimated in the form of a vector bias as described in Section 3.1.

The total measurement counts are listed in the number column of Table 3 in Section 5, where the OHM-MULs have been divided into those with dipole colatitude poleward or equatorward of  $\pm 50^\circ$ . The spatial and temporal distributions of the OHM data sets are shown in Fig. 2. The top panel shows the expected high percentage of commonality between the two sampling types (triangles). The bottom panel shows peak station participation during 1965 and 1980

with a slight decrease prior to 1984. Note here that the histograms are only recording the number of stations, including their breaks, that provide measurements within a 1 yr bin, and not the total number of measurements, in which case the OHM-MUL counts would dwarf those of the OHM-1AM.

### 2.2 Magsat

The Magsat data sets used in the CM3 model are mainly those dawn and uncorrected dusk data sets used in deriving the GSFC



**Figure 2.** OHM-1AM and OHM-MUL spatial and temporal distributions. The top panel shows observatory locations where only OHM-MUL (squares), only OHM-1AM (circles), or both (triangles) data sampling types are used in the CM3 model (cylindrical equidistant projection). The bottom panel shows a histogram of the number of OHM-MUL and OHM-1AM stations contributing data to that particular 1 yr bin across the time span of the CM3 model.

DAWN(6/83) and DUSK(6/83) models of Langel & Estes (1985a), respectively. Though these data sets are described in detail by Langel & Baldwin (1991), a brief synopsis of the processing is given here: the Magsat data were initially screened with the three-hourly  $K_p$  index by choosing data only during periods when  $K_p < 1^-$  following a period when  $K_p \leq 2^0$ . Vector data poleward of  $\pm 50^\circ$  dipole latitude were excluded to minimize the effects of FACs and ionospheric currents in the auroral regions. Scalar intensity data were retained at all latitudes. A data selection algorithm was then applied separately for both the dawn and dusk data within a  $\pm 20$  nT  $D_{st}$  level for the time intervals of 1979 November and December; 1980 January and February, and 1980 March and April. The objective of the algorithm was to obtain a uniform data distribution in both time and space. Finally, after elimination of outliers with respect to the GSFC(9/80) model of Langel *et al.* (1982), passes from slightly more disturbed times were added to sparse regions in order to improve geographic coverage. These data sets will be referred to simply as the ‘Magsat dawn’ and ‘Magsat dusk’ data sets.

The CM3 model represents a significant advancement in how fields of external origin, such as those from ionospheric coupling and EEJ currents, are parametrized. It is then desirable to include high (poleward of  $\pm 50^\circ$ ) dipole latitude vector data in the analysis, particularly the  $X$  and  $Y$  components, which are sensitive to high-latitude FACs. Accordingly,  $X$  and  $Y$  vector components were added at sampling points poleward of  $\pm 50^\circ$ , corresponding to those already providing scalar measurements in the CM3 Magsat dawn and dusk data sets. These new data sets will be referred to as the ‘Magsat polar dawn’ and ‘Magsat polar dusk’ data sets.

Before the Magsat dawn, polar dawn, dusk and polar dusk data sets were actually analysed in the CM3 model, they underwent an additional episode of refinement via the rejection of outliers with respect to a model derived in the preliminary stages of this study. This process is discussed in more detail in Section 4.2, and the resulting measurement counts for these data sets are listed in the number column of Table 3 (see Section 5).

### 2.3 POGO

The bulk of the parent POGO data sets from which those used in the CM3 model were extracted were also used to derive the POGO(2/72) field model (Langel *et al.* 1980), with additional OGO-6 data from 1969 to 1971 for magnetically quiet to moderately quiet times. These parent data sets, described in detail by Langel & Baldwin (1991), were found to have an uneven distribution in local time and to be somewhat larger than necessary for this study. Therefore, these data sets were decimated in order to achieve a more uniform geographic and magnetic local time distribution, and a more manageable size. Furthermore, the same outlier rejection phase was applied to POGO data as to Magsat data, details of which are found in Section 4.2. The resulting data sets will be referred to collectively as the ‘POGO decimated’ data set. The MLT distribution of this data set is shown in the bottom panel of Fig. 3 as a histogram of the number of measurement positions falling within 1 h MLT bins. Of the total 22 685 positions, most (1509) fall between 1:00 and 2:00 am and the least (390), unfortunately, fall between 12:00 and 1:00 pm.

Historically, the mechanism used for the POGO decimated data set did not admit entire satellite tracks. Much of the external field current system is transient and, while this distribution probably gives a broad sampling of those variations, it may not give coherent data tracks across these patterns. Because of this, it is felt that some advantage might be gained by incorporating some individual passes

of data. Accordingly, a selection of data from typical passes from quiet periods has been made. These data were also put through the outlier rejection phase, which resulted in 6754 measurements from 170 passes. This data set is referred to as the ‘POGO pass’ data set. The angular positions of the pass loci are shown in the top panel of Fig. 3, and a histogram of the number of passes that cross the geographic equator within 1 h MLT bins is shown in the middle panel. The spatial and temporal distributions appear to be sufficient to sample most of the Sq and EEJ features of the ionospheric current systems.

## 3 PARAMETRIZATION OF FIELD SOURCES

As alluded to earlier, the ultimate worth of the CM3 model lies in its ability to properly describe as much of the intended signal as possible in the data. These data, however, are limited in their spatial and temporal sampling of the field and are contaminated at some level by systematic and random error processes. An efficient model parametrization will take these limitations into account in order to achieve optimal results. The parametrizations used in this study are now described by source.

### 3.1 Core and lithospheric fields

The current systems responsible for both the core and lithospheric magnetic fields lie entirely below the regions sampled by permanent observatories and satellites. Therefore, these fields may be expressed as gradients of internal potential functions of the form

$$V_{cl}(t, \mathbf{r}) = \Re \left\{ a \sum_{n=1}^{N_{\max}} \sum_{m=0}^n \left( \frac{a}{r} \right)^{n+1} \gamma_n^m(t) Y_n^m(\theta, \phi) \right\} \quad (3)$$

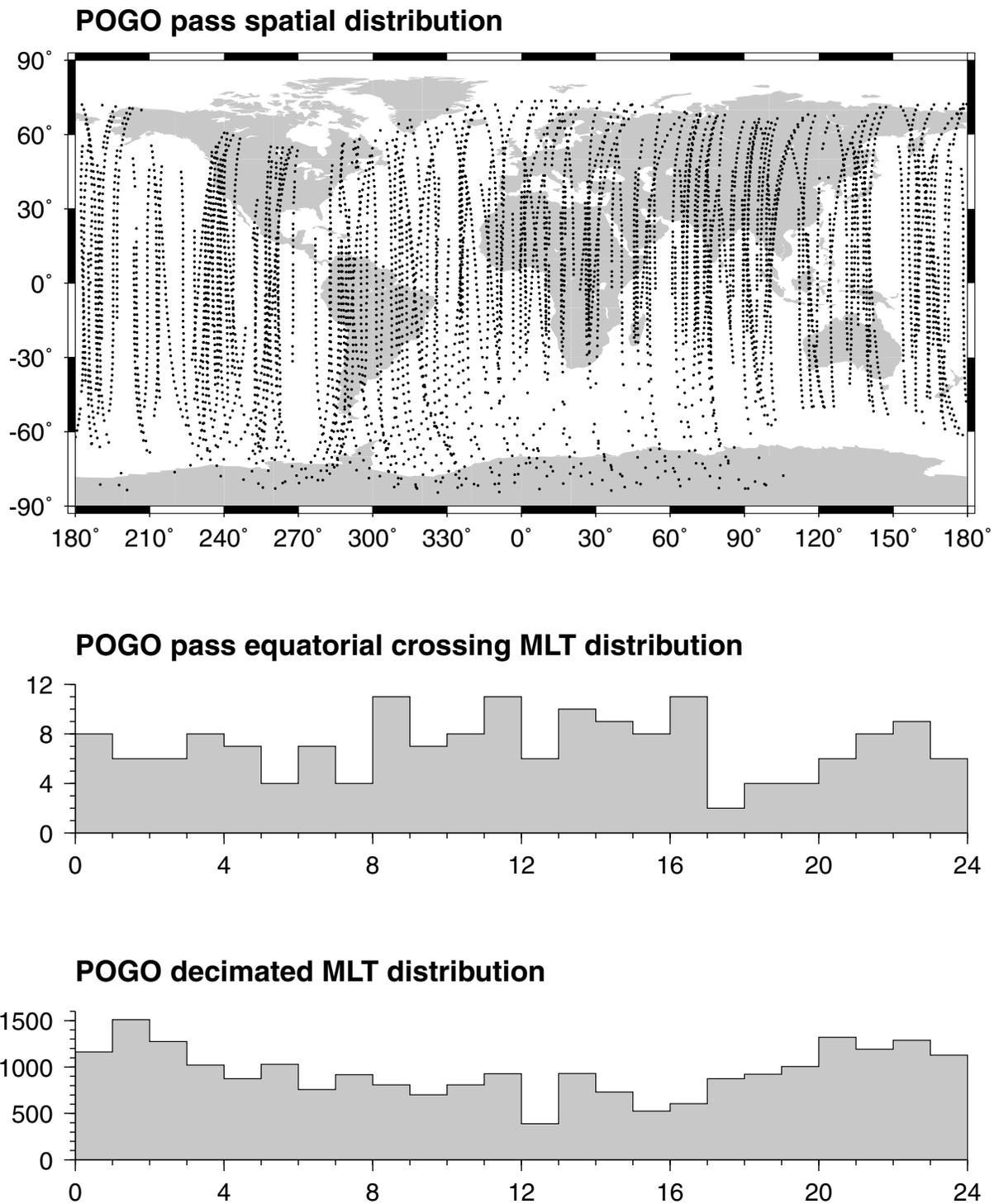
with

$$Y_n^m(\theta, \phi) = P_n^m(\cos \theta) \exp im\phi, \quad (4)$$

where  $a$  is the mean radius of the Earth (6371.2 km),  $\mathbf{r}$  is the position vector, and  $Y_n^m$  and  $P_n^m$  are the Schmidt normalized surface spherical harmonic and associated Legendre function of degree  $n$  and order  $m$ , respectively (e.g. Langel 1987). The  $\Re\{\cdot\}$  operation takes the real part of the expression only. Hence,  $\gamma_n^m$  are unique complex expansion coefficients, also known as Gauss coefficients. They are related to the usual real Gauss coefficients  $g_n^m$  and  $h_n^m$  according to  $\gamma_n^m = g_n^m - ih_n^m$ . Typically, terms in eq. (3) have been retained only up to a degree truncation level,  $N_{\max}$ , that is justified by the data, or in the case of satellites, up to the degree at which it is believed that the lithospheric field begins to dominate the series (the main field), taken by Langel & Estes (1982) to be 13. Spherical harmonic models of lithospheric fields have been derived from data with estimates of the main, magnetospheric and ionospheric fields removed. Such models indicate that noise becomes dominant somewhere between  $N_{\max} = 60$  and 70 (Ravat *et al.* 1995). That noise, however, also reflects an inaccurate estimation of the other fields. One of the intentions of the present study is to examine whether a combined model can be more effective in accurately representing the lithospheric field, and so  $N_{\max} = 65$  in this study.

#### 3.1.1 Main field secular variation

The main field is mostly dominated by contributions that originate in the core dynamo region of the Earth. The consequence of this action is that the main field varies with timescales of the order of



**Figure 3.** POGO satellite spatial and temporal distributions. The top panel shows the angular positions of the pass loci for the POGO pass data set (cylindrical equidistant projection). The middle panel shows a histogram of the number of passes from the POGO pass data set that cross the geographic equator within a particular 1 h MLT bin. The bottom panel shows a histogram of the number of measurement positions in the POGO decimated data set that fall within a particular 1 h MLT bin.

centuries. Variations with timescales shorter than a year or two are effectively screened by finite mantle conductivity, and as a result, are not detectable at the Earth's surface. Sabaka *et al.* (1997) modelled the longer timescale variations of the main field by parametrizing its SV with cubic  $B$ -splines, which are described in detail by Schumaker (1981).

This approach is adopted here using cubic  $B$ -splines and the 1980 Magsat epoch for all  $\gamma_n^m(t)$ . Because both POGO and Magsat satellite data will be included in the analysis, the time span of the model is chosen to be 1960–1985, allowing for an extension at either end of the missions. Given that the spline knot set must be defined over the entire time envelope, and choosing an equispaced knot distribution

at 2.5 yr intervals for all  $\gamma_n^m(t)$ , this results in 13 cubic  $B$ -splines per  $\gamma_n^m(t)$ . Including the offset,  $\gamma_{n0}^m$ , this gives a total of 14 parameters describing the temporal behaviour of each  $\gamma_n^m(t)$ .

Assuming that the temporal variation of terms with  $n > 13$  is negligible for the present model, the final working expression for the core and lithospheric potential is given by

$$V_{cl}(t, \mathbf{r}) = \Re \left\{ a \sum_{n=1}^{13} \sum_{m=0}^n \sum_{q=0}^{13} \left(\frac{a}{r}\right)^{n+1} \gamma_{nq}^m Y_{nq}^m(t, \theta, \phi) + a \sum_{n=14}^{65} \sum_{m=0}^n \left(\frac{a}{r}\right)^{n+1} \gamma_n^m Y_n^m(\theta, \phi) \right\}, \quad (5)$$

with

$$Y_{nq}^m(t, \theta, \phi) = \begin{cases} Y_n^m(\theta, \phi) & \text{for } q = 0, \\ Y_n^m(\theta, \phi) \int_{1980}^t b_q(\tau) d\tau & \text{for } q > 0, \end{cases} \quad (6)$$

where  $b_q(\tau)$  is the  $q$ th cubic  $B$ -spline of the expansion. The number of real coefficients in the main and lithospheric field model expansions are 2730 and 4160, respectively, which gives a total of 6890 real coefficients in this portion of the model.

### 3.1.2 Observatory biases

Permanent magnetic observatories are extremely sensitive to the short-wavelength field of the lithosphere; even at  $N_{\max} = 65$  there is undoubtedly deviation owing to an unmodelled high-degree lithospheric signal. Hence, this signal is point-sampled by the observatories and can be represented by a local vector bias,  $\mathbf{B}_{\text{bias}}$ , as introduced by Langel *et al.* (1982). There are 427 individual observatory time-series in the combined OHM-1AM and OHM-MUL data sets for which a full vector bias is estimated. Owing to data quality issues, only the horizontal data are used and bias components estimated for the Amberley II OHM-MULs and the San Fernando OHM-1AMs and OHM-MULs. Similarly, only the vertical data are used at Amatsia, Gornotayezhnaya, Norway Station and Sheshan, and hence, only this component of the biases is estimated. Note that although data from a particular station may be present in both the OHM-1AM and OHM-MUL data sets, separated biases are estimated. This has little effect, however, and was done mainly to expedite the replacement of the OAMs by the OHM-1AMs. It will be resolved in future versions of the model. The biases account for 1296 real coefficients in the model.

## 3.2 Ionospheric field

The morphology of geomagnetic variations produced by the ionospheric dynamo is relatively fixed in MLT. However, within this basic morphology there is considerable variation depending upon other phenomena such as the season, the solar cycle, the influence of the Earth's main field, etc. The ionospheric Sq currents are confined to the  $E$ -region dynamo layer (100–140 km altitude), with maximum current density at about 110 km altitude. Since the thickness of this layer (<50 km) is small compared with the distance to the point of observation (>100 km), the true ionospheric current can be well approximated by an equivalent sheet current at  $h = 110$  km altitude, which produces the same magnetic field outside the current distribution. Because these fields vary with time, there are corresponding induced currents in the Earth, with attendant fields. Since the measurements used are not acquired in the regions where the source currents flow, the fields can be represented by gradients of potential

functions. The basis functions representing the ionospheric and associated induced potentials are best understood as being built from a set of ‘elemental’ potential functions reflecting a single spatial harmonic modulated by single seasonal and diurnal periods. For the region  $a \leq r \leq a + h$  these have the form

$$V_{nsp}^m(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ l_{nsp}^m S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}) + \epsilon_{nsp}^m S_{nsp,e}^m(t, t_{\text{mut}}, \mathbf{r}) \right\}, \quad (7)$$

and for the region  $r > a + h$  these have the form

$$(V')_{nsp}^m(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ [l_{nsp}^m + (\epsilon')_{nsp}^m] S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}) \right\}, \quad (8)$$

with

$$S_{nsp,e}^m(t, t_{\text{mut}}, \mathbf{r}) = a \left(\frac{r}{a}\right)^n P_n^m(\cos \theta_d) \times \exp i(m\phi_d + \omega_s t + \omega_p p t_{\text{mut}}), \quad (9)$$

$$S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}) = a \left(\frac{a}{r}\right)^{n+1} P_n^m(\cos \theta_d) \times \exp i(m\phi_d + \omega_s t + \omega_p p t_{\text{mut}}), \quad (10)$$

where the fundamental seasonal angular frequency is  $\omega_s = 2\pi$  rad yr<sup>-1</sup> with associated wavenumber  $s$  and the fundamental diurnal angular frequency is  $\omega_p = 2\pi/24$  rad h<sup>-1</sup> with associated wavenumber  $p$ . Hence,  $\epsilon_{nsp}^m$ ,  $(\epsilon')_{nsp}^m$  and  $l_{nsp}^m$  are unique complex expansion coefficients of the external and internal ionospheric and the internal induced potentials, respectively, having a single spatial harmonic as prescribed by  $n$  and  $m$ , which oscillates on two timescales as prescribed by  $s$  and  $p$ , and propagates in one direction as prescribed by the relative signs of  $s$ ,  $p$  and  $m$ . However, the ionospheric potentials above and below the sheet source are by no means independent because the radial components of the resulting fields are continuous across the sheet. This results in the linear relationship

$$(\epsilon')_{nsp}^m = -\left(\frac{n}{n+1}\right) \left(\frac{a+h}{a}\right)^{2n+1} \epsilon_{nsp}^m, \quad (11)$$

such that eq. (8) may be rewritten in terms of  $\epsilon_{nsp}^m$ .

### 3.2.1 Induction

At this point one could define an ensemble of  $V_{nsp}^m$  as the working form of the ionospheric and associated induced potentials. However, incorporating a few conditions gained from some basic physical insights can drastically reduce the number of free parameters in this part of the model. The first insight that can be made is that  $l_{nsp}^m$  is, in general, not independent of  $\epsilon_{nsp}^m$ . The nature of this dependence is related to the conductivity structure of the crust and upper mantle, which leads to the linear relationship

$$l_{ksp}^l = \sum_{n,m} q_{knsmp}^{lm} \epsilon_{nsp}^m, \quad (12)$$

where  $q_{knsmp}^{lm}$  is an element of the complex matrix representation,  $\mathbf{Q}$ , of the transfer function between the driving ionospheric signal and the driven induced signal at similar frequencies (Schmucker 1985; Olsen 1999). For a general 3-D mantle conductivity, the  $\mathbf{Q}$  matrix is dense. However, for purposes of this model the 1-D conductivity distribution, i.e. depending only on radius, of Olsen (1998) has been adopted. It is a four-layer conductivity model derived from Sq and  $D_{\text{st}}$  data at selected European observatories. Consequently,  $\mathbf{Q}$  is now diagonal and its elements depend only upon  $n$  and the frequency  $f$ . This means that one external coefficient induces only one internal coefficient, i.e.  $l_{nn}^m(f) = q_{nn}^{mm}(f) \epsilon_n^m(f)$  (where the  $f$  dependence

now replaces the  $s$  and  $p$  indices). The mechanics of calculating  $\mathbf{Q}$  from a given conductivity model is beyond the scope of this paper, but the interested reader is referred to Parkinson & Hutton (1989).

The value of  $|q_{nn}^{mm}(f)|$  typically decreases with decreasing frequency, and therefore fields induced from higher frequencies generally dominate those from lower ones. Therefore, for simplicity, a single frequency (the highest of the diurnal–seasonal modulation) may be assigned to the matrix elements operating on  $\epsilon_{nsp}^m$ , which leads to the assignment rules

$$q_{nn}^{mm}(f) = \begin{cases} 0 & \text{for } p = 0 \text{ and } s = 0, \\ q_{nn}^{mm}(0) & \text{for } p = 0 \text{ and } |s| > 0, \\ q_{nn}^{mm}(p) & \text{for } p > 0. \end{cases} \quad (13)$$

Note that  $f=0$  is formally used to designate ‘very long-period’ rather than an absence of oscillation, in which case  $q_{nn}^{mm}(f \rightarrow 0) = 0$ . This is indeed true for the purely seasonal versus diurnal oscillations included in the model. For deriving the ‘very long-period’ response  $q_{nn}^{mm}(f \rightarrow 0)$ , the mantle is assumed to be an insulator in the region  $a - \delta \leq r \leq a$  and superconducting in  $r < a - \delta$ . For this study, a value of  $\delta = 1000$  km is used, corresponding to induction with periods longer than 1 week or so.

### 3.2.2 Quasi-dipole symmetry

A second insight is that many ionospheric phenomena are naturally organized with respect to the geometry of the Earth’s magnetic field owing to its influence on the motion of charged particles. Consequently, ionospheric conductivity is highly anisotropic, resulting in values that are so high parallel to the field that the magnetic field lines are nearly equipotential lines. Therefore, it is often convenient to work in a coordinate system that is aligned with the magnetic field, such as the QD coordinates proposed by Richmond (1995). If one could exploit the symmetries of such a coordinate system, then great savings would be realized in the number of free parameters required. For example, modelling the EEJ with spherical harmonics expanded in dipole coordinates requires both high-degree and high-order terms because of the undulation of the EEJ in  $\theta_d$ . However, the EEJ is always located at  $\theta_q = 90^\circ$ , and therefore one may be able to fit this feature with far fewer QD symmetric functions.

Before proceeding, certain properties of the QD coordinates must be articulated. First, the Laplacian operator does not separate in QD coordinates, rendering closed-form solutions out of the question. Secondly, though the QD colatitude,  $\theta_q$ , and longitude,  $\phi_q$ , do chart the sphere, their coordinate lines change slightly with radius,  $r$ . With this in mind, consider a set of basis functions, which possess the QD symmetry in two dimensions rather than three, perhaps on a constant-coordinate surface. If a sphere is chosen, then it is natural to think of the QD angular coordinates as the formal arguments to the usual surface harmonics. These functions may be expanded in terms of the surface harmonics in dipole coordinates as

$$P_k^l(\cos \theta_q) \exp i l \phi_q = \sum_{n=1}^{N_{\max}} \sum_{m=-\min(n, M_{\max})}^{\min(n, M_{\max})} (d_{kn}^{lm})^* P_n^m(\cos \theta_d) \exp i m \phi_d, \quad (14)$$

where  $\theta_q = \theta_q(r, \theta_d, \phi_d)$ ,  $\phi_q = \phi_q(r, \theta_d, \phi_d)$ ,  $d_{kn}^{-l-m} = (d_{kn}^{lm})^*$ , and the asterisk denotes complex conjugation. The regression coefficients,  $d_{kn}^{lm}$ , are determined by a standard spherical transform, where  $N_{\max}$  and  $M_{\max}$  are chosen such that sufficient convergence is achieved. The required temporal dependence is achieved by simply multiplying by the appropriate complex exponential, e.g.  $\exp i(\omega_s t + \omega_p p t_{\text{mut}})$ .

If one settles for the radial dependences found in eqs (9) and (10), then the new functions will satisfy Laplace’s equation in dipole coordinates for external and internal sources, respectively, and will exhibit the desired QD symmetry on a given sphere. Since  $h = r - a = 110$  km is the approximate height of the ionospheric  $E$ -region current system, then this should be a satisfactory choice. Hence, the potentials may be expressed in terms of

$$T_{ksp,e}^l(t, t_{\text{mut}}, \mathbf{r}) = \sum_{n,m} (d_{kn}^{lm})^* \left( \frac{a}{a+h} \right)^{n-1} S_{nsp,e}^m(t, t_{\text{mut}}, \mathbf{r}), \quad (15)$$

$$T_{ksp,i}^l(t, t_{\text{mut}}, \mathbf{r}) = \sum_{n,m} (d_{kn}^{lm})^* \left( \frac{a+h}{a} \right)^{n+2} S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}), \quad (16)$$

where the summations over  $n$  and  $m$  are understood from eq. (14). Clearly, it follows that

$$\begin{aligned} T_{ksp,e}^l \Big|_{r=a+h} &= T_{ksp,i}^l \Big|_{r=a+h} \\ &= (a+h) P_k^l(\cos \theta_q) \exp i(l\phi_q + \omega_s t + \omega_p p t_{\text{mut}}). \end{aligned} \quad (17)$$

A global plot of  $\Re\{T_{45,0,0,e}^3\}|_{r=a+h}$  is shown in Fig. 4 which illustrates the QD geometry on the sphere where the ionospheric currents are assumed to flow.

The previous development suggests a new set of ‘elemental’ potential functions, analogous to those of eqs (7) and (8), the basis of which consists of  $T_{ksp,e}^l$  and  $T_{ksp,i}^l$ , which indicates that imposing the QD symmetry at  $r = a + h$  imposes linear constraints on the original expansion coefficients. Thus, by combining eqs (11), (12), (15) and (16), one arrives at expressions for the expansion coefficients of the original ionospheric and induced potentials, given by

$$\epsilon_{nsp}^m = \sum_{k,l} d_{kn,e}^{lm} \tilde{\epsilon}_{ksp}^l, \quad (18)$$

$$(\epsilon')_{nsp}^m = \sum_{k,l} z_{kn}^{lm} \tilde{\epsilon}_{ksp}^l, \quad (19)$$

$$f_{nsp}^m = \sum_{k,l} f_{kns}^{lm} \tilde{\epsilon}_{ksp}^l, \quad (20)$$

where

$$d_{kn,e}^{lm} = \left( \frac{a}{a+h} \right)^{n-1} d_{kn}^{lm}, \quad (21)$$

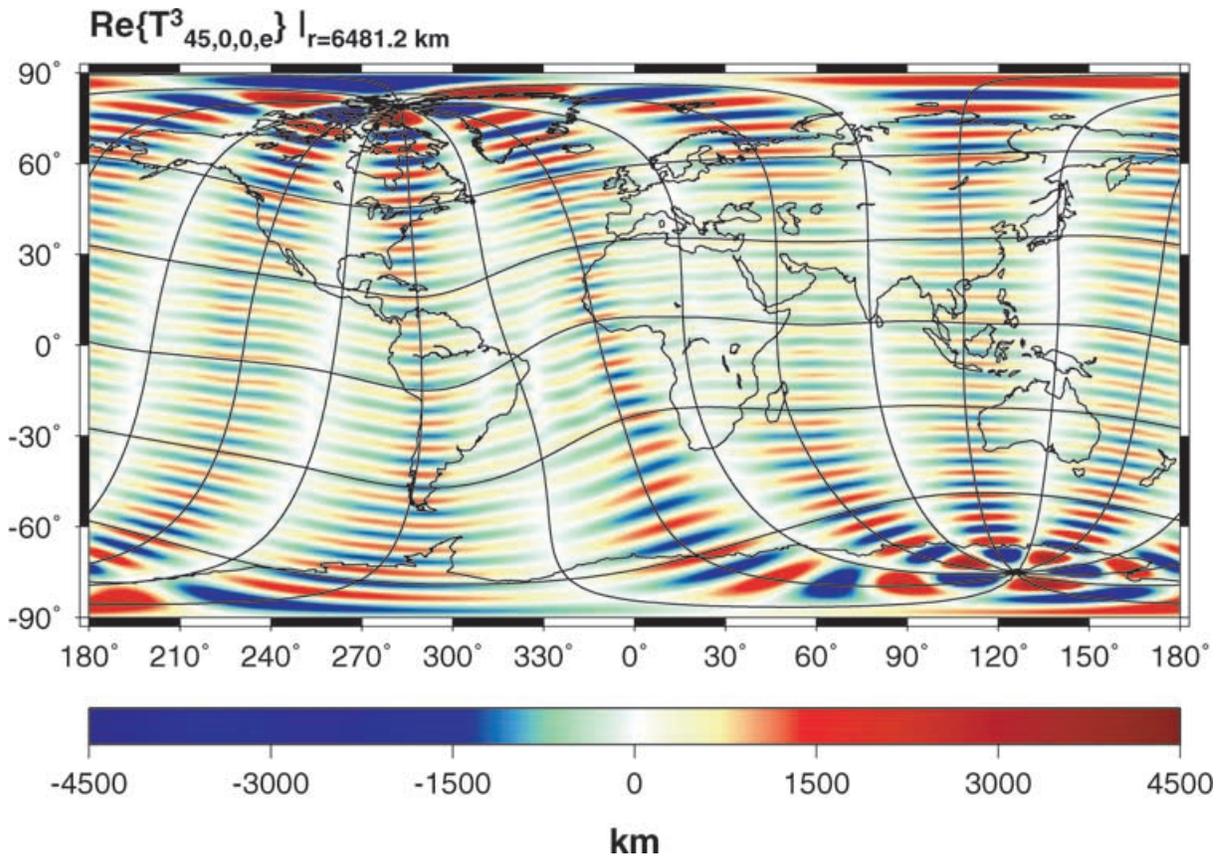
$$z_{kn}^{lm} = - \left( \frac{n}{n+1} \right) \left( \frac{a+h}{a} \right)^{n+2} d_{kn}^{lm}, \quad (22)$$

$$f_{kns}^{lm} = q_{nsp}^{mm} d_{kn,e}^{lm}, \quad (23)$$

in terms of the QD ionospheric expansion coefficients,  $\tilde{\epsilon}_{ksp}^l$ .

### 3.2.3 Solar activity

Ionospheric contributions also depend upon solar activity. In a paper on the variability of geomagnetic daily variations with solar activity, Olsen (1993) estimates the proportionality between the coefficients of a spherical harmonic analysis of the variations and sunspot number. Solar flux, however, is probably a better parameter for describing the short-term variability of solar activity of the kind encountered in this study. Therefore, the QD ionospheric expansion coefficients are redefined with a dependence upon the solar radiation flux index,  $F_{10.7}$ , as



**Figure 4.** The function  $\Re\{T^3_{45,0,0,e}\}$  evaluated on the sphere  $r = 6481.2$  km (cylindrical equidistant projection). The reference model used to define the QD coordinate system is the DGRF 1980 model (IAGA Division I Working Group 1, 1981). The  $\theta_q$  and  $\phi_q$  coordinate lines are also shown in  $30^\circ$  increments.

$$\tilde{\epsilon}_{ksp}^l = \tilde{\epsilon}_{ksp}^l (1 + NF_{10.7}). \quad (24)$$

The proportionality  $N$  is not solved for, but rather an independently estimated *a priori* value is used, which is assumed to be equal for all coefficients. Thus,  $N$  was estimated using the technique of Olsen (1993), resulting in a value of  $N = 14.85 \times 10^{-3} (10^{-22} \text{ Wm}^{-2} \text{ Hz}^{-1})^{-1}$ . This means that increasing solar flux inflates the whole ionospheric current system (and induced contributions) without changing its shape. In this study, annual averages of absolute  $F_{10.7}$  values were used for all data sets except the OHM-1AMs, which were inadvertently assigned monthly averages of adjusted  $F_{10.7}$  values. Adjusted values are corrected for the changing Sun–Earth distance, while absolute values have an additional correction for uncertainties in antenna gain and for waves reflected from the ground, which amounts to multiplying the adjusted value by 0.9. For more information on the  $F_{10.7}$  index, see <http://web.ngdc.noaa.gov/stp/SOLAR/FLUX/flux.html>. However, this is felt to be a minor problem since nightside ionospheric  $E$ -region conductivity is so low; a property that is enforced in this model (see Section 4.1.2).

### 3.2.4 Basis function selection

Given the temporal distribution of the magnetic measurements, the first four diurnal harmonics can probably be resolved. This corresponds to  $p = 1, \dots, 4$ , which is to say, the 24, 12, 8 and 6 h periods. The  $p = 0$  level will also be included for reasons to be discussed later. Instead of using all QD longitudinal wavenumbers  $l$ , it is suitable to focus on local time modes ( $p = l$ ) plus slightly faster and slightly

slower travelling modes. This limits  $l$  to a tight bandwidth,  $L$ , centred on  $p$ . In this study,  $L = 1$ . To obtain similar QD latitudinal resolution across diurnal periods, the maximum  $k$  should be a constant offset,  $K$ , from  $l$ . It is also desirable to have different resolution levels for local ( $p = l$ ) versus non-local ( $p \neq l$ ) time modes, in which case  $K$  becomes a function of  $p - l$  (i.e.  $K = K(p - l)$ ). In this study, they happen to be the same, with  $K(0) = K(\neq 0) = 40$ , chosen in hopes of resolving the EEJ. Note that this general selection scheme for  $p$ ,  $k$  and  $l$  has been used in previous studies to produce global field maps from Sq currents (e.g. Malin & Gupta 1977; Schmucker 1999).

Seasonal variation is, to first order, a function of the angle between the Earth–Sun line and the Earth’s rotation axis. However, it is also influenced by the Earth’s magnetic field, for which the dipole portion alone wobbles about the rotation axis daily. This, together with comingling of other effects, such as the solar radiation flux with its own associated periodicities, makes for a very complex picture indeed. It is believed, however, that the annual and semi-annual periods will still dominate the seasonal spectrum, and so these are the modes that will be considered in this model. Furthermore, both eastward and westward counterparts of these modes will be included. All this translates into a range of  $s = -2, \dots, 2$  for the model.

The number of columns of the matrices associated with elements  $d_{kn,e}^{lm}$ ,  $f_{kns,p}^{lm}$  and  $z_{kn}^{lm}$  are specified by the  $s$ ,  $p$ ,  $k$  and  $l$  ranges. The number of rows is determined by  $N_{\max}$  and  $M_{\max}$  in eq. (14), which are taken to be 60 and 12, respectively, based upon values of  $\max(k) = 45$  and  $\max(|l|) = 5$ . This gives a total of 1368 real regression coefficients per  $T_{ksp,e}^l$ .

Thus, the pair of expressions for the ionospheric and associated induced potentials for the regions  $a \leq r \leq a + h$  and  $r > a + h$  may be written as

$$V_{\text{ion}}(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ \sum_{s=-2}^2 \sum_{p=0}^4 \sum_{l=p-1}^{p+1} \sum_{k=\max(1,|l|)}^{|l|+40} \tilde{\epsilon}_{ksp}^l \sum_{n=1}^{60} \sum_{m=-\min(n,12)}^{\min(n,12)} \right. \\ \left. \times \left[ (d_{kn,e}^{lm})^* S_{nsp,e}^m(t, t_{\text{mut}}, \mathbf{r}) \right. \right. \\ \left. \left. + (f_{kns,p}^{lm})^* S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}) \right] \right\}, \quad (25)$$

$$V'_{\text{ion}}(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ \sum_{s=-2}^2 \sum_{p=0}^4 \sum_{l=p-1}^{p+1} \sum_{k=\max(1,|l|)}^{|l|+40} \tilde{\epsilon}_{ksp}^l \right. \\ \left. \times \sum_{n=1}^{60} \sum_{m=-\min(n,12)}^{\min(n,12)} \left[ (z_{kn}^{lm})^* + (f_{kns,p}^{lm})^* \right] \right. \\ \left. \times S_{nsp,i}^m(t, t_{\text{mut}}, \mathbf{r}) \right\}. \quad (26)$$

The quadruple summation over  $s, p, l$  and  $k$  implies 5520 real coefficients. This is about six times less than the number of parameters that would be needed if QD symmetry were not considered.

In summary, Sq currents are organized spatially and temporally by the interaction of the ambient magnetic field and the Sun. Hence, to model the associated fields efficiently this geometry is exploited in the design and selection of the basis functions. This reduces to a judicious choice of linear combinations of the ‘elemental’ functions so as to basically follow QD symmetry in space and Sun-synchronicity in time, allowing for some non-local time interaction and solar activity dependence. Proper coupling of primary and induced fields and of fields above and below the equivalent sheet current gives a representation of these fields in terms of a single set of expansion coefficients,  $\tilde{\epsilon}_{ksp}^l$ .

### 3.3 Magnetospheric field

The field of the magnetosphere is dominated by features that vary with ring-current intensity, season and solar wind parameters, but also contains features that are relatively fixed in MLT. These time-varying fields also induce currents and resultant secondary fields in the conductive portions of the crust and mantle. Because the major contributing sources of the magnetospheric field are the magnetotail, magnetopause and ring-current complexes, which lie well outside the sampling region, the field may be represented by the gradient of a potential function. In fact, the form and development of this function will closely parallel that of the ionosphere and associated induced fields for the region  $a \leq r \leq a + h$ . Therefore, one may anticipate a final working form for the magnetospheric potential very similar to eq. (25), and expedite its development by discussing only the deviations from the ionospheric case.

In addition to the regular daily and seasonal periodicities, a modulation of the magnetospheric fields with the  $D_{\text{st}}$  index, given by Langel & Estes (1985a) as

$$\epsilon_{1sp}^m = \mu_{1sp}^m + \mu_{1sp,D_{\text{st}}}^m D_{\text{st}}(t), \quad (27)$$

is used here, where the  $D_{\text{st}}$  index is in units of nT and is tabulated at hourly intervals as a function of UT (see <http://swdcd.b.kugi.kyoto-u.ac.jp/>). Note that this relationship is adopted only for the dipole terms ( $n = 1$ ), and that the temporal variability of  $D_{\text{st}}(t)$  is modulated by both seasonal and diurnal oscillations to help describe any local time asymmetries.

#### 3.3.1 Induction

Induced fields may be treated just as they were in the ionospheric field development section. The major difference is the inclusion of basis functions that are now dependent upon the seasonally and diurnally modulated  $D_{\text{st}}(t)$  index. Hence, for magnetospheric induction, eq. (13) becomes

$$q_{nn}^{mm}(f) = \begin{cases} 0 & \text{for } p = 0, \quad s = 0, \quad \text{and no } D_{\text{st}}(t), \\ q_{nn}^{mm}(0) & \text{for } p = 0, \quad s = 0, \quad \text{and } D_{\text{st}}(t), \\ q_{nn}^{mm}(0) & \text{for } p = 0 \quad \text{and } |s| > 0, \\ q_{nn}^{mm}(p) & \text{for } p > 0. \end{cases} \quad (28)$$

Magnetospheric contributions that vary only with  $D_{\text{st}}(t)$  are assumed to contain mostly signals with a period of a few days, at least during magnetically quiet days, hence, the use of a  $q_{nn}^{mm}(0)$  based upon  $\delta = 1000$  km.

#### 3.3.2 Basis function selection

At the source region for the magnetospheric current systems the Earth’s magnetic field is more dipole-like compared with ionospheric current systems, and therefore, there is no benefit in using QD coordinates to characterize magnetospheric sources. Again,  $l$  is chosen to reside in a narrow band about  $p$ , as specified by  $L$ , and maximum  $k$  is at a constant offset,  $K$ , from  $l$  to preserve latitudinal resolution levels across  $p$ .  $K$  is a function of local versus non-local time, as before. A seasonal variation of  $s = -2, \dots, 2$ , a diurnal variation of  $p = 0, \dots, 5$ , and  $L = 1$  are also used. In contrast with the ionosphere, the latitudinal resolution level, which is much less for the magnetosphere and different for local and non-local time modes, is described by  $K(0) = 5$  and  $K(\neq 0) = 3$ . It is expected that a significant number of coefficients will be negligible at these truncation levels. In an investigation of geomagnetic daily variations as predicted by the Tsyganenko model of the magnetosphere, Olsen (1996) concluded that the only non-negligible coefficients are found within the limits of  $p = 0, \dots, 2$ ,  $L = 1$ , and  $K(0) = 3$  and  $K(\neq 0) = 1$ . However, since this is the first time ionospheric and magnetospheric parametrizations of this type have been coestimated, these limits are chosen more liberally.

Thus, the magnetospheric and associated induced potentials may be written as

$$V_{\text{mag}}(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ \sum_{s=-2}^2 \sum_{p=0}^5 \sum_{l=p-1}^{p+1} \sum_{k=\max(1,|l|)}^{|l|+K(p-l)} \mu_{ksp}^l \left[ S_{ksp,e}^l(t, t_{\text{mut}}, \mathbf{r}) \right. \right. \\ \left. \left. + (q_{kks,p}^{ll})^* S_{ksp,i}^l(t, t_{\text{mut}}, \mathbf{r}) \right] \right. \\ \left. + \sum_{s=-2}^2 \sum_{p=0}^5 \sum_{l=p-1}^{p+1} \sum_{k=\max(1,|l|)}^1 \mu_{ksp,D_{\text{st}}}^l \right. \\ \left. \times D_{\text{st}} \left[ S_{ksp,e}^l(t, t_{\text{mut}}, \mathbf{r}) \right. \right. \\ \left. \left. + (q_{kks,p,D_{\text{st}}}^{ll})^* S_{ksp,i}^l(t, t_{\text{mut}}, \mathbf{r}) \right] \right\}. \quad (29)$$

Note the  $s, p$  and  $D_{\text{st}}$  indexing on the  $\mathbf{Q}$  elements, which is consistent with eq. (28). The summations entail 800 real coefficients.

### 3.4 Fields from ionospheric coupling currents

In reality, ionospheric currents do not flow in isolated shells, but are coupled to the magnetospheric and ionospheric  $E$ -region currents at the geomagnetic conjugate point through ‘coupling currents’ that

flow along the field lines of the Earth's magnetic field. This means that the Magsat sampling region between  $a+350$  and  $a+550$  km will be penetrated by  $F$ -region current for which the associated magnetic field,  $\delta\mathbf{B}$ , will not be curl-free, and hence, will not be expressible as the gradient of a potential. It should also be noted that in general only vector samples, as opposed to scalar ones, can detect these fields, since  $\delta\mathbf{B}$  is almost always perpendicular to the main field. Hence, only vector measurements from the Magsat satellite will be considered here.

Consider the toroidal–poloidal decomposition of  $\delta\mathbf{B}$ . Backus (1986) showed that for a shell with thickness that is small in comparison with its mean radius, the poloidal part goes to zero. This is the case for the Magsat sampling shell, and so  $\delta\mathbf{B}$  is considered purely toroidal such that

$$\delta\mathbf{B}(t, \mathbf{r}) = \nabla \times \mathbf{r}\Phi(t, \mathbf{r}), \quad (30)$$

with associated current density

$$\mu_0\mathbf{J}(t, \mathbf{r}) = \nabla \times \nabla \times \mathbf{r}\Phi(t, \mathbf{r}), \quad (31)$$

where  $\Phi$  is the toroidal scalar function or stream function. In order for  $\delta\mathbf{B}$  to be unique, the mean value of  $\Phi$  over the sphere must be zero. If  $\Phi$  is expanded in surface spherical harmonics as

$$\Phi(t, \mathbf{r}) = \Re \left\{ \sum_{n,m} \phi_n^m(t, r) P_n^m(\cos \theta_d) \exp im\phi_d \right\}, \quad (32)$$

then this requires  $\phi_0^0(t, r) = 0$ .

As for the radial dependence of  $\phi_n^m(t, r)$ , Olsen (1997a) considered a  $1/r$  dependence on all coefficients, which leads to a purely radial poloidal current density in eq. (31). When applying this to Magsat data, the assumption of radial currents is reasonable for the FACs in polar latitudes as well as for a very narrow band at the dip equator where the meridional current system flows in the radial direction. However, the assumption fails at middle latitudes. Nevertheless, this assumption is also adopted here, which leads to

$$\Phi(t, \mathbf{r}) = \Re \left\{ \sum_{n,m} \phi_n^m(t) S_{n,j}^m(\mathbf{r}) \right\}, \quad (33)$$

where

$$S_{n,j}^m(\mathbf{r}) = \left( \frac{a}{r} \right) P_n^m(\cos \theta_d) \exp im\phi_d. \quad (34)$$

### 3.4.1 Quasi-dipole symmetry

The  $F$ -region conductivity structure is also highly aligned with the magnetic field, which suggests the use of QD symmetry. For example, the meridional coupling currents of the EEJ show a strong radial upwelling along  $\theta_q = 90^\circ$ . This suggests functions of the form

$$T_{k,j}^l(\mathbf{r}) = \sum_{n,m} (d_{kn}^{lm})^* \left( \frac{a+h}{a} \right) S_{n,j}^m(\mathbf{r}). \quad (35)$$

It follows that

$$T_{k,j}^l \Big|_{r=a+h} = P_k^l(\cos \theta_q) \exp il\phi_q. \quad (36)$$

The  $d_{kn}^{lm}$  regression coefficients would be slightly different from those introduced earlier since these now reflect the QD symmetry at  $r = a + 450$  km. As before, the new basis imposes linear constraints on the original expansion coefficients such that

$$\phi_n^m(t) = \sum_{k,l} d_{kn,j}^{lm} \tilde{\phi}_k^l(t), \quad (37)$$

where

$$d_{kn,j}^{lm} = \left( \frac{a+h}{a} \right) d_{kn}^{lm} \quad (38)$$

and  $\tilde{\phi}_k^l(t)$  are the new QD coefficients.

### 3.4.2 Time dependence

As in the  $E$ -region, the temporal variation of the magnetic fields from the coupling currents will have strong local time modes, which are modulated by both interactions with the main field and by significant seasonal effects. As before, one could introduce the variation through factors of the form  $\exp i(\omega_s st + \omega_p pt_{\text{mut}})$ . However, only Magsat data are sensitive to these parameters, and this data has some severe limitations. First, the mission duration was approximately from 1979 November to 1980 April, about six months. If one is interested in both annual and semi-annual seasonal periods, then it is unlikely that both a phase and amplitude can be resolved. Assuming, however, that the maximum annual amplitudes occur at the solstices, then the phase may be fixed, leaving only the amplitude to be determined. If it is also assumed that maximum semi-annual amplitudes occur at the solstices and equinoxes, then its phase may also be fixed. This leads to admissible basis functions of the form

$$T_{ks,j}^l(t, \mathbf{r}) = \sum_{n,m} (d_{kn,j}^{lm})^* S_{ns,j}^m(t, \mathbf{r}), \quad (39)$$

where

$$S_{ns,j}^m(t, \mathbf{r}) = S_{n,j}^m(\mathbf{r}) \cos \omega_s st, \quad (40)$$

with associated coefficients  $\tilde{\phi}_{ks}^l$ .

Secondly, the Magsat data are sampled at only two local times, dawn and dusk. The Nyquist frequency for this sampling rate would be the diurnal ( $p = 1$ ) frequency; all higher-frequency harmonics would be aliased. Therefore, a continuous local time analysis is prohibited at the periods of interest. Consider, however, that over a period of several days to weeks a high-density distribution of both Magsat dawn and dusk passes may be obtained over all longitudes. This suggests that one model the behaviour of particular local times separately as a function of geographic position and season.

### 3.4.3 Basis function selection

Separate dawn and dusk contributions are used, each with a seasonal wavenumber range of  $s = 1, 2$ . The  $k$  and  $l$  ranges for these contributions are less complicated to specify than for the  $E$ -region, since they reflect the spatial resolution level in an Earth-fixed mode. Consequently, the latitudinal and longitudinal resolution levels are independent of one another. It is anticipated that most features of interest will lie in relatively thin, elongated QD latitude bands, which implies a high  $K_{\text{max}}$  and a low  $L_{\text{max}}$ . If the field from the EEJ coupling currents has a half-wavelength of about  $5^\circ$  at  $\theta_q = 90^\circ$ , then  $K_{\text{max}} = 40$  should suffice. For this study,  $L_{\text{max}} = 4$ . Given these values of  $K_{\text{max}}$  and  $L_{\text{max}}$ , an expansion with  $N_{\text{max}} = 60$  and  $M_{\text{max}} = 12$  is considered sufficient, i.e. a total of 1368 real regression coefficients per  $T_{ks,j}^l$ . Thus, the stream function may be expressed as

$$\Phi(t, \mathbf{r}) = \Re \left\{ \sum_{s=0}^2 \sum_{k=1}^{40} \sum_{l=0}^{\min(k,4)} \tilde{\phi}_{ks}^l \times \sum_{n=1}^{60} \sum_{m=-\min(n,12)}^{\min(n,12)} (d_{kn,j}^{lm})^* S_{ns,j}^m(t, \mathbf{r}) \right\}. \quad (41)$$

**Table 1.** Number of parameters describing each field source in the CM3 model.

Field source	Number of parameters
Observatory biases	1296
Core/lithosphere	6890
Ionosphere	5520
Magnetosphere	800
Coupling currents	2088
Total	16 594

The summations entail 1044 real coefficients, so 2088 for both dawn and dusk. This is about four times fewer than the number of parameters if QD symmetry were not considered.

### 3.5 Observed fields

The expressions for predicting the magnetic field as seen by Magsat, POGO and OHMs are given by

$$\mathbf{B}_{\text{Magsat}} = -\nabla_{\text{g}} V_{\text{cl}} - \mathbf{R}_{\text{gd}}[\nabla_{\text{d}}(V'_{\text{ion}} + V_{\text{mag}}) - \nabla_{\text{d}} \times \mathbf{r}\Phi], \quad (42)$$

$$\|\mathbf{B}_{\text{Magsat}}\| = \|-\nabla_{\text{g}} V_{\text{cl}} - \mathbf{R}_{\text{gd}}\nabla_{\text{d}}(V'_{\text{ion}} + V_{\text{mag}})\|, \quad (43)$$

$$\|\mathbf{B}_{\text{POGO}}\| = \|-\nabla_{\text{g}} V_{\text{cl}} - \mathbf{R}_{\text{gd}}\nabla_{\text{d}}(V'_{\text{ion}} + V_{\text{mag}})\|, \quad (44)$$

$$\mathbf{B}_{\text{OHM}} = -\mathbf{R}_{\text{es}}[\nabla_{\text{g}} V_{\text{cl}} + \mathbf{R}_{\text{gd}}\nabla_{\text{d}}(V'_{\text{ion}} + V_{\text{mag}})] + \mathbf{B}_{\text{bias}}, \quad (45)$$

where  $V_{\text{cl}}$ ,  $V_{\text{ion}}$ ,  $V'_{\text{ion}}$  and  $V_{\text{mag}}$  are the potential functions for the fields of the core and lithosphere, the ionospheric  $E$ -region for  $a \leq r \leq a + h$  and  $r > a + h$ , and the magnetosphere, respectively.  $\nabla_{\text{g}}$  and  $\nabla_{\text{d}}$  are the gradient operators in geographic and dipole spherical polar coordinates, respectively.  $\mathbf{R}_{\text{gd}}$  is a rotation matrix from the local dipole spherical to the local geographic spherical basis.  $\mathbf{R}_{\text{es}}$  is a rotation matrix from the local geographic spherical to the local geographic ellipsoidal basis, where the IAU ellipsoid (International Astronomical Union 1966) is used (Barton 1997). A summary of the number of parameters describing each field source is given in Table 1.

## 4 ESTIMATION OF MODEL PARAMETERS

Having essentially defined the forward problem for this study in eqs (42)–(45), the model parameters were found by solving the inverse problem via the weighted least-squares method by minimizing the cost function

$$\mathcal{L}(\mathbf{x}) = [\mathbf{d} - \mathbf{a}(\mathbf{x})]^T \mathbf{W}(\mathbf{d} - \mathbf{a}(\mathbf{x})) + \sum_{i=1}^k \lambda_i (\mathbf{x} - \mathbf{x}_{a_i})^T \Lambda_i (\mathbf{x} - \mathbf{x}_{a_i}) \quad (46)$$

with respect to the model parameter vector  $\mathbf{x}$ , where  $\mathbf{d}$  is the measurement vector,  $\mathbf{a}(\mathbf{x})$  is the model prediction vector (eqs 42–45),  $\mathbf{x}_{a_i}$  is the preferred parameter vector from the  $i$ th constraint system,  $\mathbf{W}$  and  $\Lambda_i$  are the uncalibrated inverse covariance matrices of the data noise vector and the deviation vector from  $\mathbf{x}_{a_i}$ , respectively,  $\lambda_i$  is the damping parameter from the  $i$ th constraint system, and  $k$  is the number of constraint systems. The Gauss iterative method was used to estimate the optimal model. Its  $n$ th step is given by (Tarantola & Valette 1982)

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \left[ \mathbf{A}_n^T \mathbf{W} \mathbf{A}_n + \sum_{i=1}^k \lambda_i \Lambda_i \right]^{-1} \times \left[ \mathbf{A}_n^T \mathbf{W} (\mathbf{d} - \mathbf{a}(\mathbf{x}_n)) + \sum_{i=1}^k \lambda_i \Lambda_i (\mathbf{x}_{a_i} - \mathbf{x}_n) \right], \quad (47)$$

where  $\mathbf{x}_{n+1}$  is the estimated model parameter vector after  $n$  steps and  $\mathbf{A}_n$  is the Jacobian of  $\mathbf{a}(\mathbf{x})$  at  $\mathbf{x}_n$ .

The uncalibrated error-covariance matrix of the final parameter estimate,  $\tilde{\mathbf{x}}$ , is

$$\mathbf{C}_{\tilde{\mathbf{x}}} = \left( \mathbf{A}^T \mathbf{W} \mathbf{A} + \sum_{i=1}^k \lambda_i \Lambda_i \right)^{-1}, \quad (48)$$

where  $\mathbf{A}$  is the Jacobian at  $\tilde{\mathbf{x}}$ . An unbiased estimate,  $s^2$ , of the data misfit is given by (Toutenburg 1982)

$$s^2 = \frac{[\mathbf{d} - \mathbf{a}(\tilde{\mathbf{x}})]^T \mathbf{W} (\mathbf{d} - \mathbf{a}(\tilde{\mathbf{x}}))}{N - \text{tr}[\mathbf{R}]}, \quad (49)$$

where

$$\mathbf{R} = \mathbf{C}_{\tilde{\mathbf{x}}} \mathbf{A}^T \mathbf{W} \mathbf{A}, \quad (50)$$

and where  $\text{tr}[\cdot]$  is the trace operator and  $N$  is the number of observations. The resolution matrix  $\mathbf{R}$  acts as a filter through which the true model state is seen and the trace of which gives the number of model parameters resolved by the data (Langel 1987). Thus,  $s^2$  is a measure of how well the model fits the weighted data per degree of freedom (DOF), and should be approximately unity if the weighting is correct (Bloxham *et al.* 1989). One may now obtain an unbiased estimate of the calibrated parameter error-covariance matrix by simply multiplying  $\mathbf{C}_{\tilde{\mathbf{x}}}$  by  $s^2$ . In this sense a ‘calibrated’ error-covariance matrix reflects how well the observations can actually be fit.

### 4.1 Regularization and *a priori* information

All of the constraints considered in this study are quadratic norms in  $\mathbf{x}$ , where  $\Lambda_i$  is the associated norm matrix, and these will now be discussed.

#### 4.1.1 Smooth main field secular variation

In order for CM3 to remain consistent throughout its temporal domain the main field SV must behave in a plausible manner. This is a critical issue because of the restricted distribution of the data sets: Magsat and POGO provide small-scale spatial information, but only for short durations, while OHMs cover the entire domain, but only provide broad-scale spatial information. Thus, for CM3 the class of admissible SV models is restricted by imposing temporal and spatial smoothing (least-structure) constraints. Since the notion of smoothing usually implies a minimization of curvature, an attempt was made to smooth the second-order time and space derivatives of the radial field SV,  $\dot{B}_r$ , through time over the CMB. The spatial derivatives actually take the form of the surface Laplacian. However, it was found that the second-order time derivative of  $\dot{B}_r$  was not restrictive enough, especially near the boundaries of the domain, and so the first-order derivative of  $\dot{B}_r$  was used instead, i.e.  $\ddot{B}_r$ .

Formally, these norms, denoted as  $\mathcal{Q}_{|\dot{B}_r|}$  and  $\mathcal{Q}_{|\nabla_{\text{g}}^2 \dot{B}_r|}$ , measure the weighted mean-square magnitude of  $\dot{B}_r$  and the surface Laplacian of  $\dot{B}_r$  over the CMB,  $\Omega$ , from 1960 to 1985, respectively. This leads to

$$\mathcal{Q}_{|\dot{B}_r|}(\mathbf{x}) = \int_{1960}^{1985} \int_{\Omega} [\dot{B}_r(t, r, \theta, \phi)]^2 d\Omega dt \times \left( \int_{1960}^{1985} \int_{\Omega} d\Omega dt \right)^{-1}, \quad (51)$$

$$\mathcal{Q}_{|\nabla_s^2 \dot{B}_r|}(\mathbf{x}) = \int_{1960}^{1985} \int_{\Omega} [\nabla_s^2 \dot{B}_r(t, r, \theta, \phi)]^2 d\Omega dt \times \left( \int_{1960}^{1985} \int_{\Omega} d\Omega dt \right)^{-1}, \quad (52)$$

where  $B_r$  and its derivatives are computed from the potential of eq. (5).

#### 4.1.2 Minimal nightside ionospheric E-region currents

It is known that the nightside ionospheric E-region conductivity is greatly diminished owing to the lack of solar extreme-ultraviolet (EUV) ionizing radiation, at least at the mid and low latitudes. This means that the equivalent sheet current density,  $\mathbf{J}_{\text{eq}}$  (which is assumed to be equal to the true ionospheric current density) is minimal in these areas. This  $\mathbf{J}_{\text{eq}}$  is a toroidal field residing on the sphere where the sheet currents flow and may be expressed in terms of a current or stream function,  $\Psi$  (Chapman & Bartels 1940), such that

$$\mathbf{J}_{\text{eq}}(t, t_{\text{mut}}, \mathbf{r}) = -\hat{\mathbf{r}} \times \nabla \Psi(t, t_{\text{mut}}, \mathbf{r}), \quad (53)$$

where

$$\Psi(t, t_{\text{mut}}, \mathbf{r}) = \Re \left\{ -\frac{1}{\mu_0} \sum_{s=-2}^2 \sum_{p=0}^4 \sum_{l=p-1}^{p+1} \sum_{k=\max(1, |l|)}^{|l|+40} \tilde{\epsilon}_{ksp}^l \times \sum_{n=1}^{60} \sum_{m=-\min(n, 12)}^{\min(n, 12)} (d_{kn,e}^{lm})^* \left( \frac{2n+1}{n+1} \right) \times S_{nsp,e}^m(t, t_{\text{mut}}, \mathbf{r}) \right\}, \quad (54)$$

and  $\mathbf{r}$  is the position vector on the sphere of radius  $r = a + h$ .

The theory takes the form of a quadratic norm,  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$ , which measures the mean-square magnitude of  $\mathbf{J}_{\text{eq}}$  on a spherical sector,  $\Omega_s$ , fixed in dipole MLT longitude (defined as  $\phi_{\text{mlt}} = \phi_d + \omega_p t_{\text{mut}}$ ) over UT. The low-conductivity sector is 8 h wide and centred on local 1:00. This leads to

$$\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}(\mathbf{x}) = \int_0^T \int_{\Omega_s} \|\mathbf{J}_{\text{eq}}(\phi_{\text{mlt}}, t_{\text{mut}}, \theta_d)\|^2 d\Omega_s dt_{\text{mut}} \times \left( \int_0^T \int_{\Omega_s} d\Omega_s dt_{\text{mut}} \right)^{-1}. \quad (55)$$

The  $t_{\text{mut}}$  integration is facilitated by two assumptions: first, the time variation of the  $F_{10.7}$  index is neglected, rendering  $\mathbf{J}_{\text{eq}}$  periodic over 1 yr such that  $T = 1$  yr may be used; and secondly, universal time  $t$  in eq. (9) is treated as  $t_{\text{mut}}$ , which is then used for time integration. Although the second assumption is best at low and mid latitudes, it is expected that this will make little difference in the analysis, especially since  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$  is a soft bound.

The  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$  norm works in conjunction with the  $p = 0$  terms of eq. (25) to establish a nightside baseline such that  $\mathbf{J}_{\text{eq}}$  is minimized at those hours. This baseline is a global function, able to adjust to geographic shifts, which is static on diurnal timescales, but varies with season. Because there is difficulty in separating this functional behaviour from that of main field SV at satellite altitude, the strength of the norm is adjusted via the associated  $\lambda_{\|\mathbf{J}_{\text{eq}}\|}$  such that all  $p = 0$

terms are determined by the norm. It is also for this reason that the influence of the norm cannot be greatly reduced in the polar regions (e.g. via some dipole colatitude weighting function) where  $\mathbf{J}_{\text{eq}}$  is thought to flow at all MLTs.

#### 4.1.3 Smooth ionospheric E-region currents

Recall that in order to resolve the EEJ along the dip equator, QD degrees of up to  $k = 45$  are used. Since the  $T_{ksp,e}^l$  functions are global, it is expected that spurious oscillations will be exhibited in the  $\mathbf{J}_{\text{eq}}$  morphology. Although the preferred model state for the  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$  norm is zero, its influence is limited in proximity to MLTs of 21–5 h. Hence, an additional norm is sought to minimize this roughness on the dayside, which suggests minimizing the mean-square magnitude of some function of the second-order horizontal derivatives of  $\mathbf{J}_{\text{eq}}$ . A natural choice is the surface Laplacian of  $\mathbf{J}_{\text{eq}}$ . However, this norm should not interfere with the  $p = 0$  baseline established for  $\mathbf{J}_{\text{eq}}$  by the  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$  norm, and so it is restricted to current densities in the  $p > 0$  regime, denoted as  $\mathbf{J}_{\text{eq}, p > 0}$ . Consequently, the norm may be applied at all MLTs. Furthermore, it must not interfere with legitimate EEJ variations near the dip equator nor with flow in the auroral regions. This may be accomplished by introducing a non-negative weighting function in  $\theta_d$ , which is smaller in the equatorial and polar regions and larger at mid-latitudes; for this study  $\sin^8 2\theta_d$  is used. Although a more rigorous approach would use QD colatitude, it is much more complicated and is left for future work.

Formally, this norm, denoted as  $\mathcal{Q}_{\|\nabla_s^2 \mathbf{J}_{\text{eq}, p > 0}\|}$ , measures the weighted mean-square magnitude of the surface Laplacian of  $\mathbf{J}_{\text{eq}, p > 0}$  on a sphere,  $\Omega$ , over UT. This leads to

$$\mathcal{Q}_{\|\nabla_s^2 \mathbf{J}_{\text{eq}, p > 0}\|}(\mathbf{x}) = \int_0^T \int_{\Omega} \|\nabla_s^2 \mathbf{J}_{\text{eq}, p > 0}(t_{\text{mut}}, \theta_d, \phi_d)\|^2 \times \sin^8 2\theta_d d\Omega dt_{\text{mut}} \times \left( \int_0^T \int_{\Omega} \sin^8 2\theta_d d\Omega dt_{\text{mut}} \right)^{-1}. \quad (56)$$

The same assumptions regarding the  $t_{\text{mut}}$  integration are made here as well. Note that the surface Laplacian operator multiplies  $S_{nsp,e}^m$  by a factor of  $-n(n+1)$ , and so  $\mathcal{Q}_{\|\nabla_s^2 \mathbf{J}_{\text{eq}, p > 0}\|}$  damps the higher-degree harmonics much more severely than  $\mathcal{Q}_{\|\mathbf{J}_{\text{eq}}\|}$ , as intended.

#### 4.1.4 Minimal magnetospheric local-time dipole deviations

It is anticipated that the magnetospheric field expansion of eq. (29) includes many more coefficients than can be reliably estimated from the data at hand, especially those describing deviations from a dipole in MLT. Experience from earlier phases of modelling suggests that excessive cross-talk or correlations between the ionospheric and non- $D_{\text{st}}$ -dependent magnetospheric expansions will probably exist as a result of poor field separation caused by limited satellite vector data coverage in local time. Therefore, a magnetospheric solution is sought which is smooth in some sense that will reduce this coupling. Specifically, define a fifth quadratic norm,  $\mathcal{Q}_{\|\Delta \mathbf{B}_{\text{ld}}\|}$ , which measures the mean-square magnitude of the deviations from a dipole in MLT ( $k > 1$  or  $l \neq p$ ) and independent of  $D_{\text{st}}$  on a sphere at Magsat altitude ( $r = a + h_m$  with  $h_m = 450$  km),  $\Omega$ , over UT. This leads to

$$\mathcal{Q}_{\|\Delta \mathbf{B}_{\text{ld}}\|}(\mathbf{x}) = \int_0^T \int_{\Omega} \|\Delta \mathbf{B}_{\text{ld}}(t_{\text{mut}}, \theta_d, \phi_d)\|^2 d\Omega dt_{\text{mut}} \times \left( \int_0^T \int_{\Omega} d\Omega dt_{\text{mut}} \right)^{-1}, \quad (57)$$

where  $\Delta \mathbf{B}_{\text{td}}$  includes all  $\mu_{ksp}^l$  terms in eq. (29) for which  $k > 1$  or  $l \neq p$ .

#### 4.1.5 Minimal ionospheric coupling currents

Recall that the radial component of the meridional coupling currents of the EEJ are being accounted for in the Magsat observations via  $\delta \mathbf{B}$  in eq. (30), requiring QD degrees of up to  $k = 40$ . Hence,  $T_{ks,j}^l$  will be susceptible to instabilities similar to those of  $T_{ksp,e}^l$  and consequently, the associated  $J_r$  for both dawn and dusk will need to be smoothed. Consider that the inclination of the Magsat orbit was such that no data were acquired within a cap of half-angle of about  $7^\circ$  centred on the geographic poles (Langel & Estes 1985a). This, combined with the fact that  $J_r$  is expressed in dipole coordinates, makes damping the polar regions a necessity for both dawn and dusk. Furthermore, since  $J_r$  shows little structure at low and mid-latitudes during dawn, there is no need to introduce a  $\theta_d$  influence function as in  $\mathcal{Q}_{\|\nabla_s^2 \mathbf{J}_{\text{eq},p>0\|}$ . Although the EEJ coupling currents are present at dusk along the dip equator, the inclusion of an influence function that is small only at low dipole latitudes is complicated, and so it is omitted in this study.

To this end, define a sixth type of quadratic norm,  $\mathcal{Q}_{|J_r|}$ , which measures the mean-square magnitude of  $J_r$  on a sphere at Magsat altitude ( $r = a + h_m$  with  $h_m = 450$  km),  $\Omega$ , over UT, given by

$$\mathcal{Q}_{|J_r|}(\mathbf{x}) = \int_0^T \int_{\Omega} [J_r(t, \theta_d, \phi_d)]^2 d\Omega dt \left( \int_0^T \int_{\Omega} d\Omega dt \right)^{-1}, \quad (58)$$

where

$$J_r(t, \mathbf{r}) = \Re \left\{ \frac{1}{\mu_0 (a + h_m)} \sum_{s=0}^2 \sum_{k=1}^{40} \sum_{l=0}^{\min(k,4)} \tilde{\phi}_{ks}^l \times \sum_{n=1}^{60} \sum_{m=-\min(n,12)}^{\min(n,12)} (a_{kn,j}^m)^* n(n+1) S_{ns,j}^m(t, \mathbf{r}) \right\}, \quad (59)$$

and  $\mathbf{r}$  is the position vector on the sphere of radius  $r = a + h_m$ . Because  $J_r$  has a period of 1 yr,  $T = 1$  yr. Again, there are separate  $\mathcal{Q}_{|J_r|}$  included for both dawn and dusk local times.

#### 4.2 Weighting

In this section the  $\mathbf{W}$  matrix in eq. (46) will be defined. The components of the data noise vector are assumed to reflect stationary, uncorrelated processes, which renders a diagonal  $\mathbf{W}$ . In particular, the scalar noise process is considered Gaussian, which is only true if the process is isotropic.

The data subsets of this study, considered to represent distinct stochastic populations, are listed as subheaders under the component column of Table 3 (see Section 5). Further assumptions are made concerning the variation of the error processes with respect to orientation: the OHM-1AM and low/mid dipole latitude ( $|90^\circ - \theta_d| < 50^\circ$ ) OHM-MUL processes are considered to be isotropic; the high dipole latitude ( $|90^\circ - \theta_d| \geq 50^\circ$ ) OHM-MUL and Magsat dawn and dusk processes are considered to be isotropic in the  $XY$ -plane; and the Magsat dawn and dusk mid/low dipole latitude processes are considered to be isotropic in the  $XZ$ -plane.

Assignment of the actual variances is accomplished by iteratively adjusting a starting set of variances similar to  $s^2$  of eq. (48), but corresponding to each data subset, until they reach unity. This requires that a model be fitted to the data. Because this is a computationally

intensive venture, a preliminary model, denoted A, developed during the initial stages of this study is used once for all the data. A detailed description of this model is provided by Sabaka *et al.* (2000). The *a priori* data noise variances resulting from this procedure are listed in the  $\sigma_v$  column of Table 3 (see Section 5).

This procedure also provides an opportunity to reject gross outliers with respect to the A model. Specifically, residuals in the Magsat mid/low and high dipole latitude data sets greater than 25 and 100 nT, respectively, are rejected as are OHM-MUL residuals greater than 150 nT and POGO residuals greater than 25 nT. The resulting measurement counts for the various data sets are listed in the number column of Table 3 (see Section 5).

#### 4.3 Application

Two iterations of the Gauss method were required from an  $\mathbf{x}_0$  provided by a preliminary model known as B, a successor to A, which is described more fully by Sabaka *et al.* (2000). This B model, also referred to as the GSFC/CU(12/96) model, has been used by Purucker *et al.* (1997) to study north–south-trending anomalies of lithospheric origin, particularly over Australia.

In the present study, the number of observations ( $\dim \mathbf{d}$ ) is 591 432, the number of model parameters ( $\dim \mathbf{x}$ ) is 16 594 and the number of quadratic smoothing norms is seven ( $\mathbf{x}_a = \mathbf{0}$  for each).

Because seven norms are employed in this study and because the number of parameters is large, a typical trade-off study between the various terms in the least-squares objective function is impractical and so the more qualitative method of visual inspection is used. This is thought to suffice for two reasons: first, the norms measure physically meaningful quantities about which qualitative knowledge abounds and which can be checked visually; and secondly,  $\lambda_i$  have a logarithmic influence on the norms and so only an order of magnitude accuracy is needed. The additional criterion of low correlation between ionospheric and magnetospheric coefficients was instrumental in selecting the  $\mathcal{Q}_{\|\Delta \mathbf{B}_{\text{td}}\|}$  damping parameter value. The values selected for  $\lambda_i$  are listed in Table 2.

### 5 RESULTS

The CM3 model may be examined for self-consistency and for physical plausibility, the first of which will now be discussed.

#### 5.1 Residuals and data fits

In this section, the data residuals will be examined. Table 3 lists, among other things, the unweighted residual statistics, i.e. the mean,  $\mu_r$ , and standard deviation about the mean,  $\sigma_r$ , of the CM3 model for each field measure for each of the data subsets of interest. Also listed is the *a priori* standard deviation of the associated data noise,  $\sigma_v$ ,  $\mu_r$

**Table 2.** Damping parameter values used in the CM3 model.

Norm	Damping parameter ( $\lambda$ )	
$\mathcal{Q}_{ \tilde{B}_r }$	$3.1 \times 10^1$	$(\text{nT yr}^{-2})^{-2}$
$\mathcal{Q}_{ \nabla_s^2 \tilde{B}_r }$	$3.1 \times 10^{-7}$	$(\text{nT yr}^{-1} \text{ km}^{-2})^{-2}$
$\mathcal{Q}_{\ \mathbf{J}_{\text{eq}}\ }$	$8.4 \times 10^2$	$(\text{A km}^{-1})^{-2}$
$\mathcal{Q}_{\ \nabla_s^2 \mathbf{J}_{\text{eq},p>0\ }$	$3.8 \times 10^{-2}$	$(\text{A km}^{-3})^{-2}$
$\mathcal{Q}_{\ \Delta \mathbf{B}_{\text{td}}\ }$	$6.3 \times 10^4$	$(\text{nT})^{-2}$
$\mathcal{Q}_{ J_r }$ dawn	$1.3 \times 10^{-5}$	$(\text{nA m}^{-2})^{-2}$
$\mathcal{Q}_{ J_r }$ dusk	$1.3 \times 10^{-5}$	$(\text{nA m}^{-2})^{-2}$

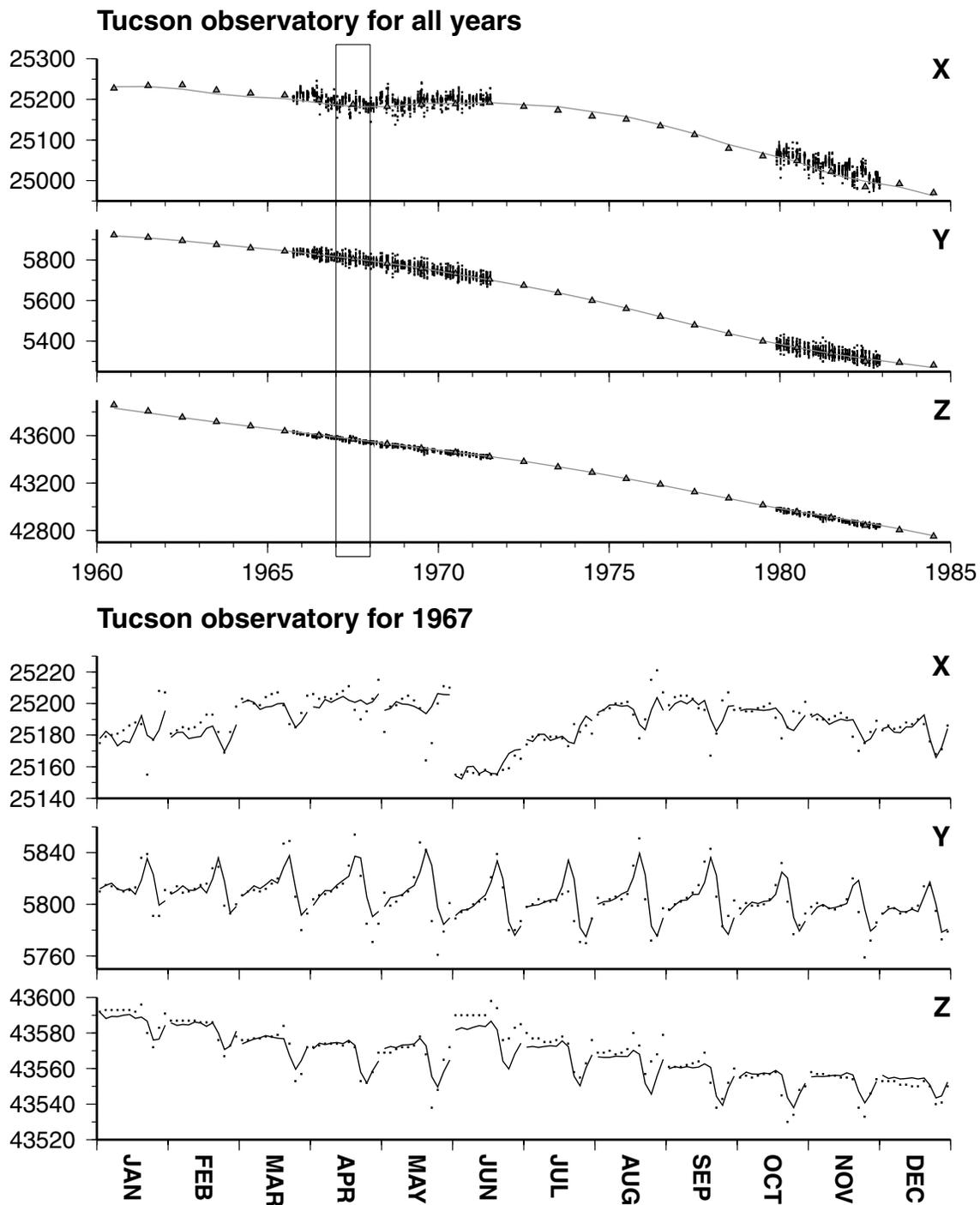
**Table 3.** Comparison of unweighted residual statistics ( $\mu_r$ ,  $\sigma_r$ , and  $\sigma_v$  in units of nT).

Component	CM3				GSFC/CU(12/96)			GSFC(8/95-SqM)		
	Number	$\mu_r$	$\sigma_r$	$\sigma_v$	Number	$\mu_r$	$\sigma_r$	Number	$\mu_r$	$\sigma_r$
<b>OAM</b>										
<i>X</i>	–	–	–	–	4047	0.0	28.5	4048	0.0	26.0
<i>Y</i>	–	–	–	–	4047	0.0	40.1	4048	0.0	26.0
<i>Z</i>	–	–	–	–	4047	0.0	33.7	4048	0.0	31.9
<b>OHM-1AM</b>										
<i>X</i>	27 094	0.0	15.0	15.0	–	–	–	–	–	–
<i>Y</i>	27 106	0.0	13.1	15.0	–	–	–	–	–	–
<i>Z</i>	27 512	0.0	19.6	15.0	–	–	–	–	–	–
<b>OHM-MUL</b>										
$ 90^\circ - \theta_d  < 50^\circ$										
<i>X</i>	54 792	0.0	10.0	11.0	56 963	0.0	11.4	153 544	0.0	14.8
<i>Y</i>	54 846	0.0	9.0	11.0	57 016	0.0	12.1	153 544	0.0	14.2
<i>Z</i>	56 696	0.0	8.5	11.0	55 978	0.0	10.2	153 544	0.0	11.1
$ 90^\circ - \theta_d  \geq 50^\circ$										
<i>X</i>	65 834	0.0	17.0	18.0	65 451	0.0	17.9	74 712	0.0	19.4
<i>Y</i>	65 895	0.0	16.5	18.0	65 487	0.0	17.6	74 712	0.0	17.5
<i>Z</i>	65 693	0.0	23.2	21.0	65 230	0.0	20.8	74 712	0.0	22.4
<b>Magsat corrected dusk</b>										
<i>X</i>	–	–	–	–	–	–	–	11 060	0.3	10.3
<i>Y</i>	–	–	–	–	–	–	–	11 060	0.3	13.4
<i>Z</i>	–	–	–	–	–	–	–	11 060	0.3	8.2
<i>B</i>	–	–	–	–	–	–	–	12 399	–0.3	7.5
<b>Magsat dusk</b>										
<i>X</i>	9381	–0.1	4.6	5.4	9381	–0.04	5.3	–	–	–
<i>Y</i>	9321	–0.03	5.7	6.8	9321	0.0	6.8	–	–	–
<i>Z</i>	9382	0.1	4.0	5.4	9382	0.1	5.4	–	–	–
<i>B</i>	11 404	0.5	3.6	5.5	11 404	1.04	5.5	–	–	–
<b>Magsat polar dusk</b>										
<i>X</i>	7985	–1.6	15.8	18.5	7985	–2.5	17.6	–	–	–
<i>Y</i>	7988	0.3	16.5	18.5	7988	1.0	17.6	–	–	–
<b>Magsat dawn</b>										
<i>X</i>	10 570	0.05	4.3	5.0	10 570	–0.3	5.1	10 595	–0.4	7.1
<i>Y</i>	10 537	–0.0005	4.4	5.4	10 537	–0.1	5.2	10 595	–0.1	7.4
<i>Z</i>	10 588	–2.0	3.4	5.0	10 588	–0.8	4.7	10 595	–0.2	6.2
<i>B</i>	12 441	–0.2	3.6	5.3	12 441	–0.6	5.2	12 460	–0.6	7.4
<b>Magsat polar dawn</b>										
<i>X</i>	8483	–0.4	17.4	19.0	8483	–1.3	18.3	–	–	–
<i>Y</i>	8445	0.1	18.2	19.0	8445	–1.8	19.4	–	–	–
<b>POGO original</b>										
<i>B</i>	–	–	–	–	–	–	–	57 434	–2.0	8.0
<b>POGO decimated</b>										
<i>B</i>	22 685	–0.2	4.2	4.8	22 685	0.0	5.1	–	–	–
<b>POGO pass</b>										
<i>B</i>	6754	0.4	5.0	5.8	6754	0.0	6.4	–	–	–

indicate whether a bias exists in the description of the data by the model. With the exception of Magsat polar dusk *X* and Magsat dawn *Z*,  $\mu_r$  have magnitudes well below 1 nT; in fact those of the OHM-1AMs and OHM-MULs are zero, by virtue of their vector biases. As expected, the mid/low dipole latitude subsets of Magsat and the OHM-MULs are fitted substantially better than their high-latitude counterparts. Within the mid/low dipole latitude Magsat dawn and dusk subsets scalar *B* is well fitted (unaffected by attitude errors) as is *Z* (least affected by external contributions) whilst *Y* is fit slightly worse (possibly caused by attitude error (R. Holme, priv. comm., 2000) or the result of dynamic variability in the meridional current system, especially at dusk). The POGO *B* are fitted slightly worse than Magsat *B*, which may be caused by their extended distribution

in time and the OHM-1AMs fits are consistent with those of the bulk OHM-MULs.

The OAMs and OHM-MULs for Tucson are shown in Fig. 5 at two different timescales, along with the values predicted by CM3. The top panel shows all available OAM data from 1960 to 1985 along with the OHM-MULs. The OAMs were predicted using annually integrated values from the magnetosphere and ionosphere. The predictions appear to be quite good for all components, especially when considering that the OAMs were not used when deriving the model. The bottom panel shows the observed and predicted OHM-MULs for the quietest day of each month in 1967. The fits are judged to be good, especially when considering the adjustment for the jump discontinuities in *X* and *Z* between the quietest days of May and June,



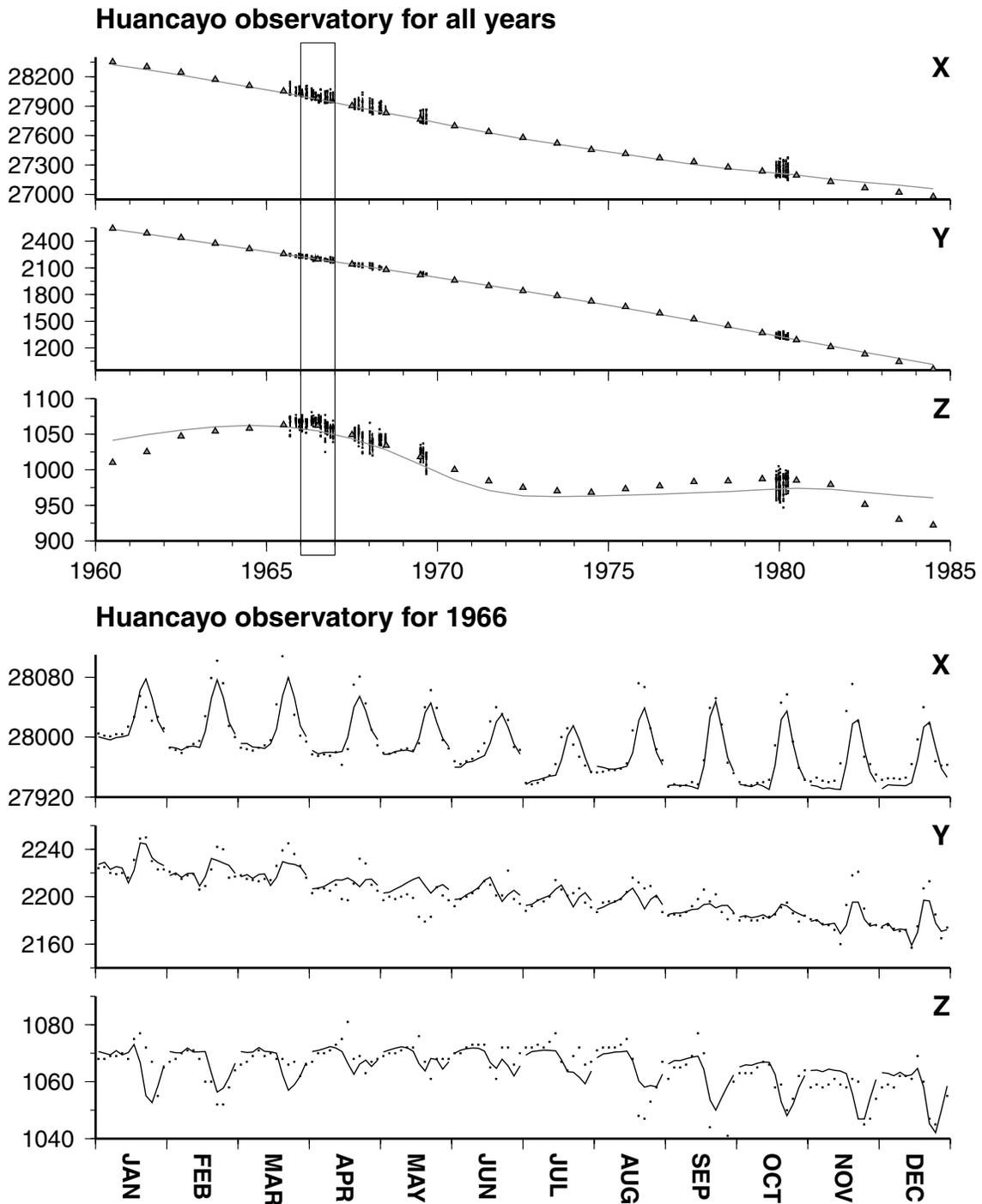
**Figure 5.** Fits of the CM3 model to the OAM and OHM-MUL vector components of the Tucson observatory. The top panel shows all measured OAMs (triangles) and OHM-MULs (squares) and their predicted values (grey line), all in nT, over the time span of the model, in years. The bottom panel shows all OHM-MULs (squares) and the model predictions (black line), all in nT, for the year 1967 (indicated by the box outline in the top panel). The abscissa in the bottom panel is discontinuous, being comprised of the quietest day of each month over the year, and begins at 00 UT for each day.

which are caused by a change in external field strength (ring-current level).

Similar plots are provided in Fig. 6 for the Huancayo observatory, which is located under the EEJ. The top panel shows a similar behaviour as in the case of Tucson, except that the Z predictions are excessively smooth, especially near the end points. This may be caused by the smoothing of the main field SV in Z, which is dominating the small-Z component at this dip equatorial station. The

bottom panel is now for the year 1966 and shows good fits to all three components. Note the adjustment in X for the jump discontinuity between the quietest days of June and July caused by different levels of ring-current activity.

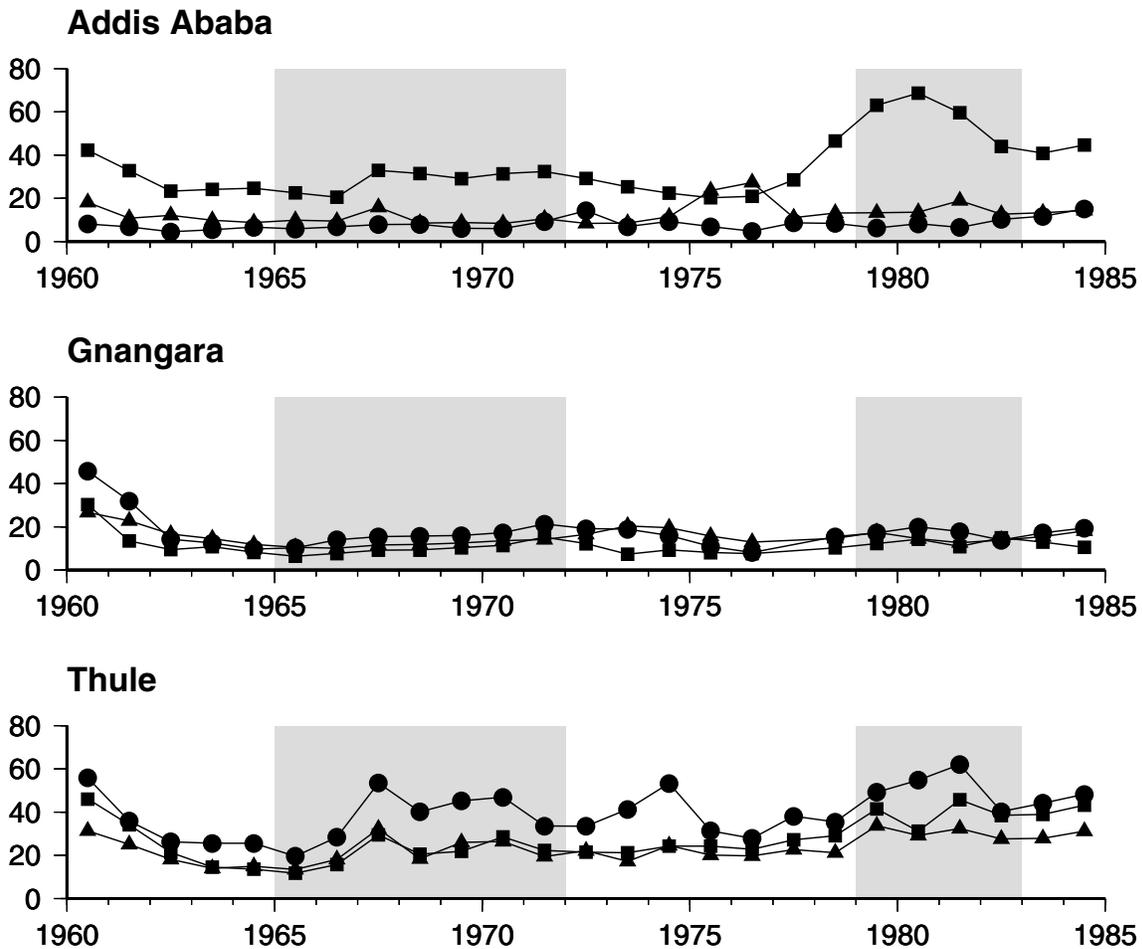
While the Tucson and Huancayo plots show that the model satisfactorily predicts the analysed OHM-MUL data in confined envelopes and the integrated effects of OAMs across the domain, it is also important that the model be able to extrapolate its full



**Figure 6.** Fits of the CM3 model to the OAM and OHM-MUL vector components of the Huancayo observatory. The top panel shows all measured OAMs (triangles) and OHM-MULs (squares) and their predicted values (grey line), all in nT, over the time span of the model, in years. The bottom panel shows all OHM-MULs (squares) and the model predictions (black line), all in nT, for the year 1966 (indicated by the box outline in the top panel). The abscissa in the bottom panel is discontinuous, being comprised of the quietest day of each month over the year, and begins at 00 UT for each day.

resolving power to poor control areas. To this end, the rms fits to the OHM-MULs from three observatories, Addis Ababa (equatorial), Gwangara (mid-latitude, southern hemisphere) and Thule (high latitude), were computed in 1 yr bins from 1960 to 1985. The results are shown in Fig. 7, with the rms *X*, *Y* and *Z* values indicated by squares, triangles and circles, respectively. The data segments within the POGO and Magsat coverage envelopes were analysed in the model and these are indicated with a shaded backdrop. All

values ramp up near 1960 and most ramp up slightly near 1985. However, in the satellite gap between 1972 and 1979, only the *Y* component of Addis Ababs shows any excursion more severe than at any other time. In fact, it could be argued that several components are actually better fitted in this gap than during the satellite coverage envelopes. Note that Addis Ababa *X* and Thule *Z* show the highest variance, as would be expected from their geographic positions.



**Figure 7.** The rms fits to OHM-MULs from the Addis Ababa (equatorial), Gngangara (mid-latitude, southern hemisphere) and Thule (high-latitude) observatories, computed in 1 yr bins from 1960 to 1985. The rms  $X$ ,  $Y$  and  $Z$  values are indicated by squares, triangles and circles, respectively. The shaded areas delineate the POGO and Magsat coverage envelopes where the OHM-MULs were actually analysed in the model. All ordinates are in nT.

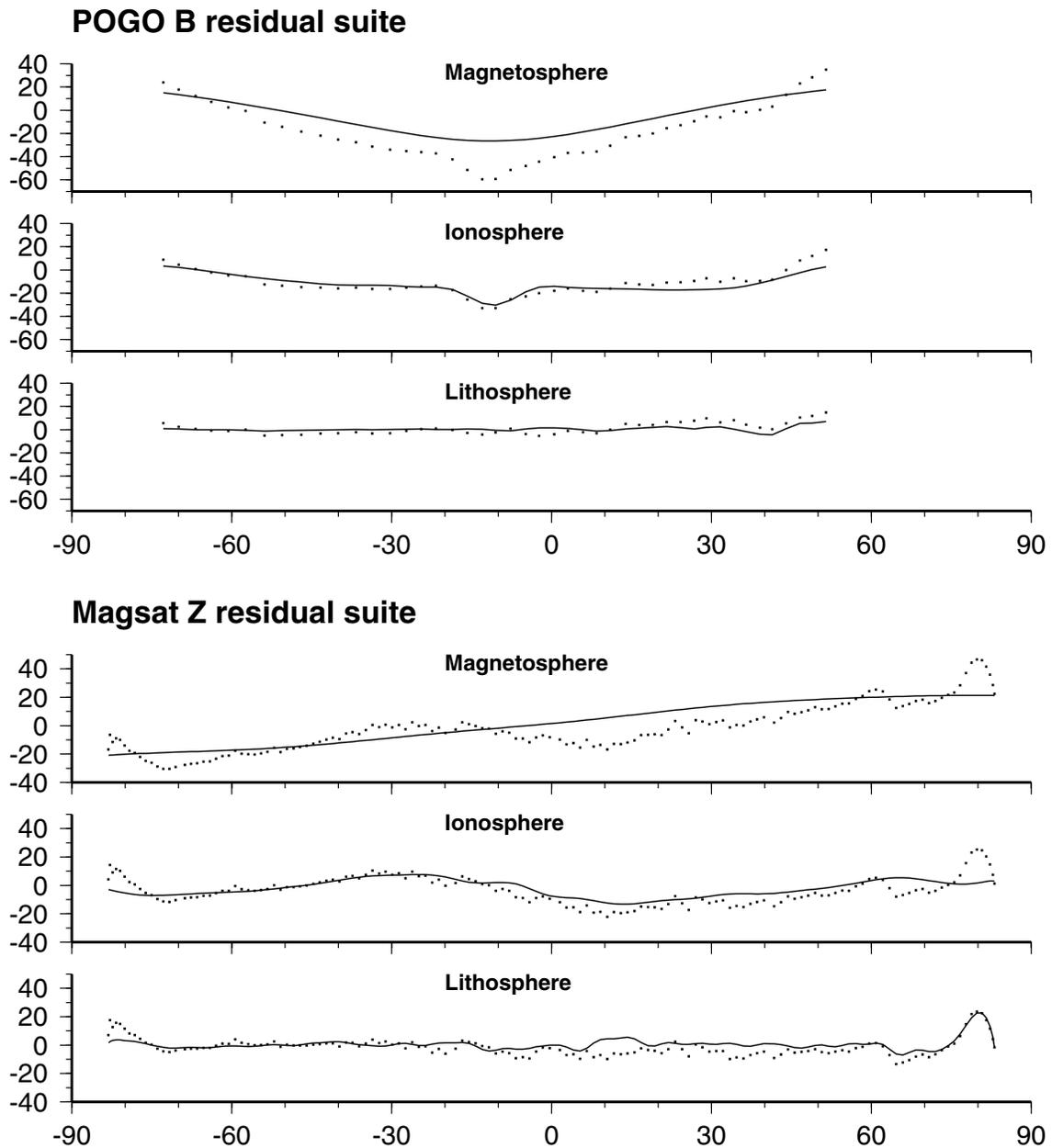
The top panel of Fig. 8 shows the fit to  $B$  for a particular pass of POGO data that was included in the model. This pass crosses the geographic equator at  $59^\circ\text{W}$  at noon MLT. This panel shows a progression of residuals from top to bottom, with a given member showing residuals,  $\Delta B$  (squares), with respect to the main field plus all preceding labelled fields,  $\mathbf{B}_{\text{cum}}$ , and the component (black line) of the predicted, currently labelled field,  $\mathbf{B}_{\text{new}}$ , in the direction of  $\mathbf{B}_{\text{cum}}$ . Note the clear EEJ signature in the residuals of the second member, which the model is able to reproduce. Though the fit is satisfactory for most of the pass ( $\leq 8$  nT), it begins to steadily diverge north of about  $20^\circ$  geographic latitude, reaching a maximum magnitude of 7.7 nT at  $51^\circ$ . This may be a result of the influence of external current systems, particularly unmodelled ring-current dynamics.

The fit to the  $Z$  component for Magsat dusk pass 263, which was not included in the model, is shown in a similar format in the bottom panel of Fig. 8. Again, the total fit is satisfactory for practically all of the pass ( $\leq 15$  nT). The corresponding  $X$ - and  $Y$ -component residual progressions are shown in Fig. 9. The fits are within about 15 nT for all of  $X$  and most of  $Y$ . The obvious exception is found in the polar region south of about  $-60^\circ$  geographic latitude. The ‘coupling currents’ member of the  $Y$  suite shows significant toroidal field presence at low latitude, which is caused by radial meridional currents connected with the EEJ impinging the Magsat

sampling shell. This model (probably the first to include the global, non-potential contribution at satellite altitude) is able to fit this feature.

## 5.2 Parameter uncertainty and separability

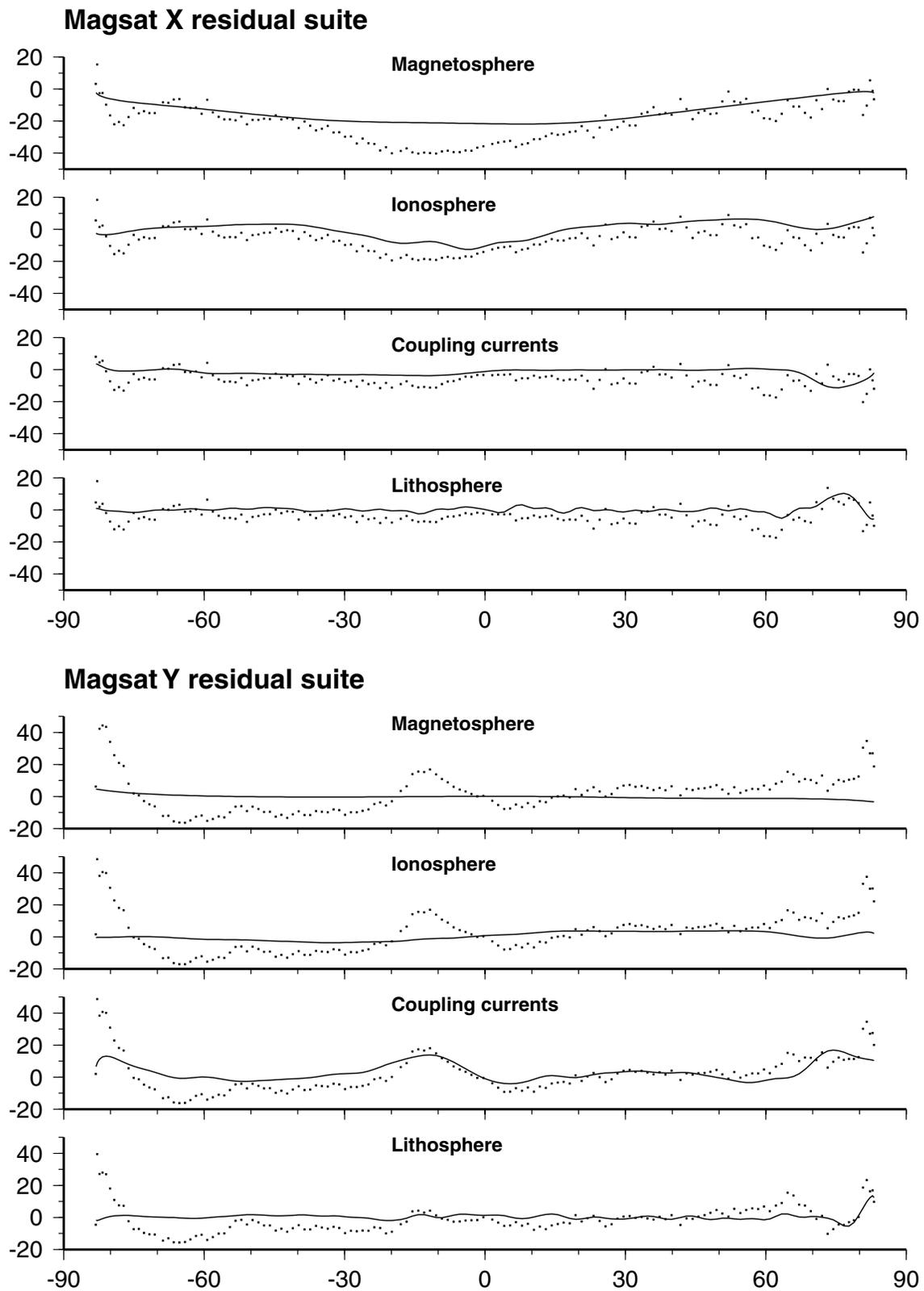
As with any inverse problem, one must address the issues of parameter uncertainty and separability. These ultimately reflect upon the quality and distribution of data and any additional regularizing information from which the parameters are inferred. It must be stressed that in this study the measures of uncertainty and separability are statistical quantities derived from a linear covariance analysis, which is applicable here since the non-linearities owing to scalar data are very slight and noise distributions are *assumed* to be Gaussian. The *a posteriori* parameter error-variances may be taken directly from the diagonals of the calibrated error-covariance matrix,  $\mathbf{C}_x$ , introduced in Section 4. However, if significant off-diagonal elements exist, then the true resolving power of the data may be obfuscated and will most likely lie along linear combinations of the parameters. A principal component analysis, or eigenvalue decomposition, of  $\mathbf{C}_x$  will, in fact, reveal these combinations and give a much clearer view of what is known or unknown by the data. Fig. 10 shows the square-root spectrum (principal standard deviations) of  $\mathbf{C}_x$ , which may be read directly as the uncertainties (in nT) of the



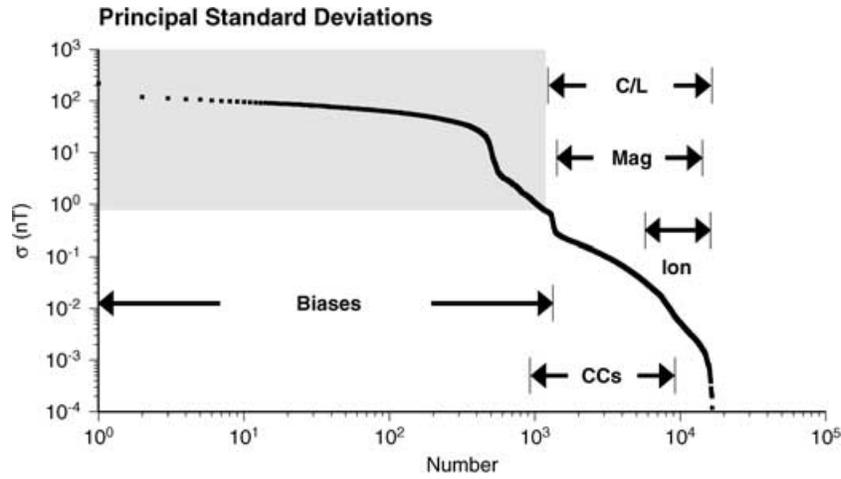
**Figure 8.** Fit of the CM3 model to the scalar ( $B$ ) values of a particular POGO pass used to derive the model and to the  $Z$  component of the Magsat dusk pass 263, which was not used in deriving the model. The POGO pass crosses the geographic equator at  $59^\circ\text{W}$  at noon while the Magsat pass crosses at  $129^\circ\text{W}$ . The top panel is a suite of residual plots for the POGO  $B$  data as a function of latitude. The progression is from the top to bottom member, with a given member showing residuals,  $\Delta B$  (squares), with respect to the main field plus all preceding labelled fields,  $\mathbf{B}_{\text{cum}}$ , and the component (black line) of the predicted, currently labelled field,  $\mathbf{B}_{\text{new}}$ , in the direction of  $\mathbf{B}_{\text{cum}}$ . The bottom panel is a similar suite of residual plots for the Magsat  $Z$  data, except the black line now represents the  $Z$  component of  $\mathbf{B}_{\text{new}}$ . All ordinates are in nT.

corresponding eigenvectors, which represent these linear combinations. Arrows indicate boundaries in which OHM bias (Biases), coupling-current (CCs), core/lithospheric (C/L), magnetospheric (Mag), or ionospheric (Ion) parameter directions account for at least 1 per cent of some of the eigenvector lengths, and so, give a sense as to how well these *groups* are known. First, the shaded area indicates the region in which 99.99 per cent of the total model error-variance is concentrated; it encompasses 1189 DOFs, and combinations outside of this region are known to better than 0.77 nT. This region is dominated almost exclusively by uncertainties in OHM biases. In particular, the least-known combinations of parameters involve several

Antarctic stations, with the South Pole  $Z$  biases having the highest uncertainty at 225 nT. There is a subregion in which eigenvectors span both biases and coupling-current parameters, indicating the presence of a slight correlation between these two non-potential fields. The eigenvectors spanning the parameters of the remaining source fields are all known very well. However, this does not mean that the individual parameters themselves are well separated. In fact, this would translate to perfect separability only if the eigenvalues (principal variances) were equal for all of these spanning eigenvectors; Fig. 10 clearly shows that this is not the case. One must look to the correlation matrix associated with  $\mathbf{C}_x$  to address separability.



**Figure 9.** Fits of the CM3 model to the  $X$  and  $Y$  components of the Magsat dusk pass 263. The top and bottom panels are suites of residual plots for the  $X$  and  $Y$  data, respectively, from this pass as a function of latitude. The progression is from the top to bottom member, with a given member showing residuals (squares) with respect to the main field plus all preceding labelled fields and the  $X$  and  $Y$  components (black lines) of the predicted currently labelled field. Note that coupling currents are now included. All ordinates are in nT.



**Figure 10.** Principal standard deviations from the calibrated error-covariance matrix,  $\mathbf{C}_x$ , of CM3. The shaded area indicates the region in which 99.99 per cent of the total model error-variance is concentrated. Arrows indicate boundaries in which OHM bias (Biases), coupling-current (CCs), core/lithospheric (C/L), magnetospheric (Mag) or ionospheric (Ion) parameter directions account for at least 1 per cent of some of the eigenvector lengths.

If the  $(ij)$ th element of the parameter error-covariance matrix is given by

$$(\mathbf{C}_x)_{ij} = \nabla_{\mathbf{a}}(\tilde{x}_i - x_i) \cdot \nabla_{\mathbf{a}}(\tilde{x}_j - x_j), \quad (60)$$

where  $\nabla_{\mathbf{a}}(\tilde{x}_i - x_i)$  is the gradient of the  $i$ th parameter error,  $\tilde{x}_i - x_i$ , with respect to the data model,  $\mathbf{a}$ , then the  $(ij)$ th element of the parameter correlation matrix,  $\mathbf{R}_x$ , is given by

$$(\mathbf{R}_x)_{ij} = \frac{\nabla_{\mathbf{a}}(\tilde{x}_i - x_i) \cdot \nabla_{\mathbf{a}}(\tilde{x}_j - x_j)}{\|\nabla_{\mathbf{a}}(\tilde{x}_i - x_i)\| \|\nabla_{\mathbf{a}}(\tilde{x}_j - x_j)\|} = \cos \varphi_{ij}, \quad (61)$$

where  $\varphi_{ij}$  is the angle between the gradients of  $\tilde{x}_i - x_i$  and  $\tilde{x}_j - x_j$ . In this study, parameter separability is defined as the degree of independence of one parameter error from another with respect to data perturbations. Indeed, if two parameters are perfectly correlated or anticorrelated, then there does not exist a data perturbation that affects one but not the other. This leads to a singular  $\mathbf{C}_x$  matrix. Conversely, if two parameters enjoy some degree of separation, then there exists a data perturbation, which affects one to some degree but not the other. Thus, the correlation matrix gives a direct statistical measure of this degree of separability.

Elements of the CM3 correlation matrix for which the absolute values exceeded 0.7 were examined for evidence of parameter separability problems. This threshold was chosen since it corresponds to  $\varphi_{ij} \approx 45^\circ$ ; geometrically, the mid-point between being fully correlated and uncorrelated.

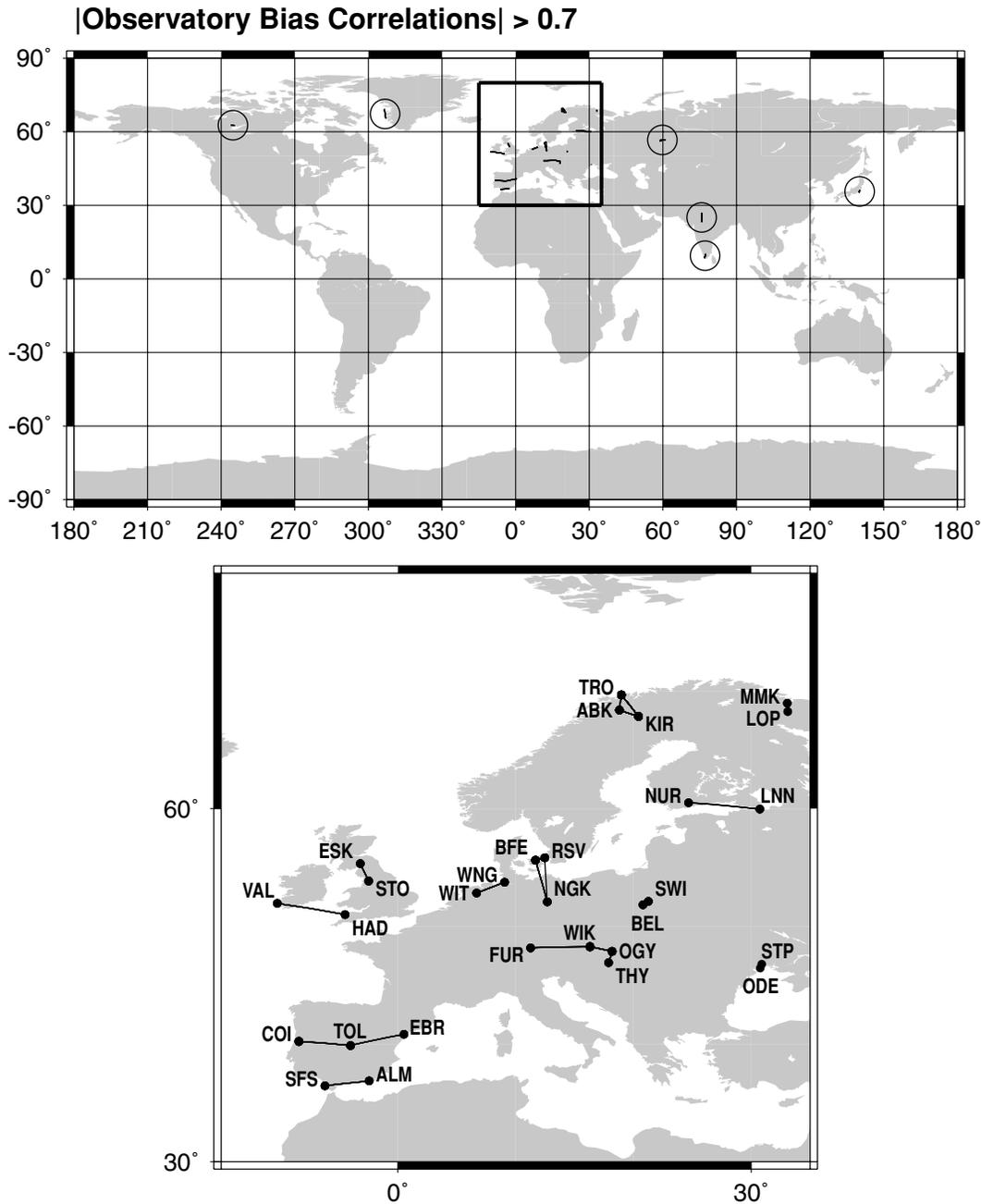
Three of the five major categories of correlations at this level involve spatial and temporal separability problems within the same field source and indicate overparametrizations: in the main field SV, where  $B$ -spline density per harmonic is high; in the ionospheric  $E$ -region, where the latitudinal resolution is high; and in the ionospheric  $F$ -region coupling currents, particularly the seasonal variation.

A fourth category includes correlations amongst OHM biases, at the same and different locations. Correlations across breaks in a particular station record may be caused by both segments cross-talking with the main field SV. Correlations between biases at spatially distinct points are intriguing because they may provide insight into spatial correlation lengths of crustal field sources on local scales. To investigate this, a global map was produced in which the locations of spatially distinct observatories were connected by a line if any bias component of one was correlated with any bias com-

ponent of the other, with an absolute value above 0.7 (see the top panel of Fig. 11). The box outlines the European sector (exploded in the bottom panel) and the circles have been added to help guide the eye to some of the more obscure lineations. First, note that none of the line segments is longer than about  $6^\circ$  of arc and most are much shorter. This corresponds well with the resolution limit of the  $N_{\max} = 65$  truncation level of the core and lithospheric field expansion, and one would expect that any coherence at lengths longer than this would be described by the model. Secondly, note that the largest concentration and length of lines are found in Europe. This follows from the fact that the highest concentration of observatories is found in Europe. If a significant crustal correlation length is present elsewhere in the world, but is only sampled by one observatory, then this exercise will fail to detect it. Although a detailed interpretation will not be attempted here, it must be mentioned that some of the lineation patterns in Europe do appear to be related to known geological features. The polygon defined by the Tromsø (TRO), Abisko (ABK) and Kiruna (KIR) observatories agrees well with the location of the Kiruna crustal anomaly, while the east–west lineation between Nurmijärvi (NUR) and Voyeykovo (LNN) may be related to the Kursk anomaly. The complex of Fürstenfeldbruck (FUR), Wien Kobenzl (WIK), O Gyalla Peth (OGY) and Tihany (THY) seems to follow the fabric of the Alpine region, while the east–west lineations in the Iberian peninsula may also be following regional trends.

The last category represents the only significant cross-talk between different field sources. It involves negative correlations between the parameters of the non- $D_{st}$ -dependent magnetosphere and the ionospheric  $E$ -region with similar spatial and temporal behaviour. In a preliminary CM3-type model that did not include the  $Q_{\parallel \Delta B_{\text{ind}}}$  smoothing, these correlations exist between harmonics having broad-scale latitudinal variation. For CM3, these correlations exceeded the 0.7 threshold only for the ‘tilt’ component ( $k = l = p = 1$ ) of the MLT dipole, and most of that is in the noon–midnight direction (real part).

While a direct inspection of  $\mathbf{R}_x$  is certainly a worthwhile exercise, it can be misleading: most physically meaningful quantities are functions of many parameters, and even though correlations between any given pair may be negligible, the cumulative affect may be significant. It is natural then to wonder whether any further cross-talk between field sources does exist, and so global correlation



**Figure 11.** Observatory bias correlations between spatially distinct locations with absolute values above 0.7 shown as connecting lines (cylindrical equidistant projection). This includes correlations between any component of one observatory bias with any component of another. Circles have been added to the top panel as a visual aid in locating some of the more obscure lineations.

maps between like pairs of vector field components from suspected sources were generated. These include magnetosphere with ionosphere (Fig. 12) and lithosphere with ionosphere (Fig. 13). Correlations between fields from coupling-currents and potential fields and correlations between lithospheric and magnetospheric fields were found to be negligible. Since it is of interest to gauge the ‘worst-case’ separability problems, these maps are generated at the summer solstice when Magsat data, probably the most dominant separating agent for the pairs of interest, are absent. Of the two pairs of field sources, the correlations between magnetosphere and ionosphere are stronger, yet the largest correlation magnitude is only 0.41, which translates to  $\varphi \approx 66^\circ$ . Therefore, field separability in CM3 is judged

to be satisfactory for most applications. Fig. 12 shows ranges of  $(-0.16, 0.21)$ ,  $(-0.02, 0.41)$ ,  $(-0.29, 0.09)$  for  $XX$ ,  $YY$  and  $ZZ$ , respectively. Since the maps are centred on noon MLT, one can see that the most dominant features appear to be Earth-fixed. This may indicate a slight cross-talk between the magnetosphere and the static portion of the ionosphere introduced as a baseline for nightside current minimization. It is also interesting to notice that on average  $YY$  is strongly positive,  $ZZ$  is strongly negative and  $XX$  is mixed.

For the lithospheric and ionospheric correlations, Fig. 13 shows ranges of  $(-0.16, 0.06)$ ,  $(-0.26, 0.06)$ ,  $(-0.29, 0.04)$  for  $XX$ ,  $YY$  and  $ZZ$ , respectively, which are predominantly negative. Note that the colour scale has now changed from Fig. 12 in order to accentuate

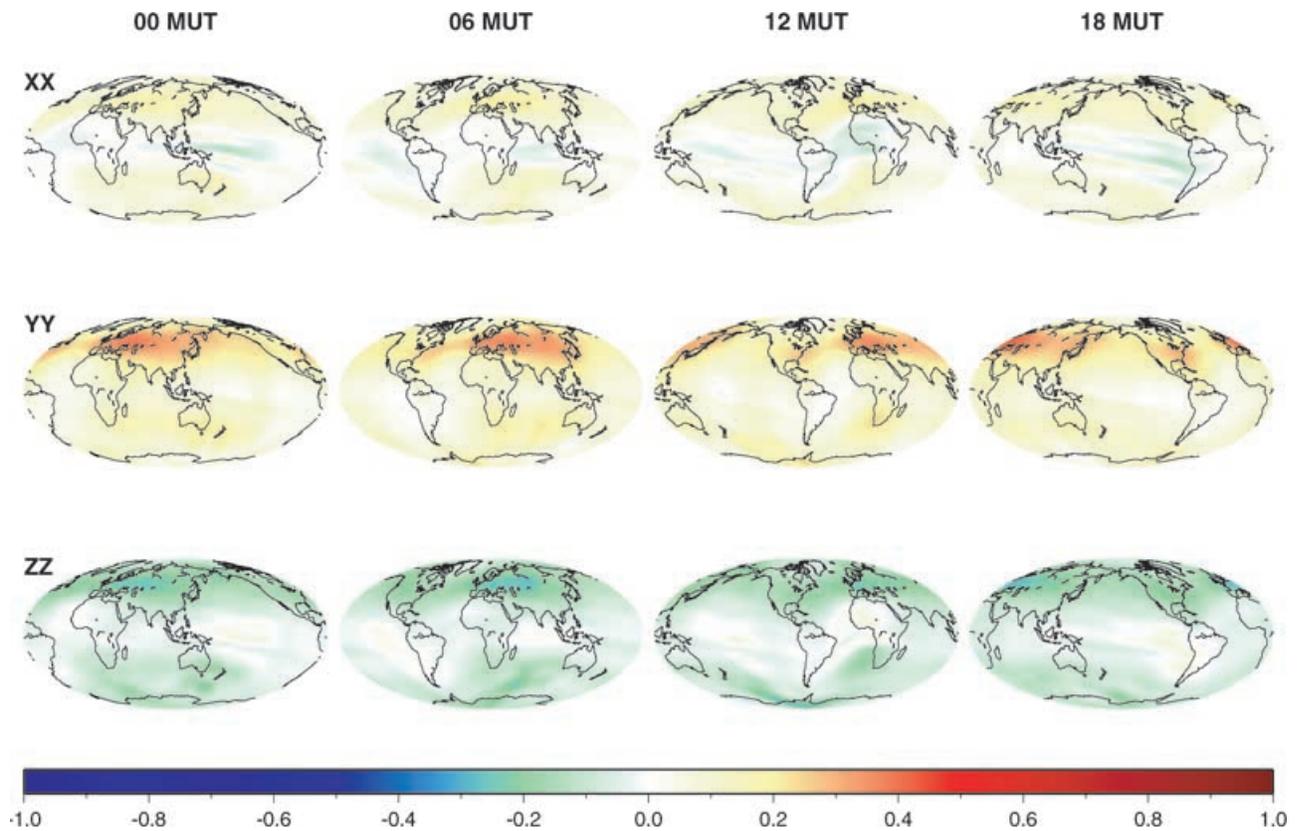


Figure 12. Correlation maps of like vector components from magnetospheric and ionospheric fields at  $r = 6821.2$  km on 1980 June 21 centred on noon MLT, but for different MUTs (Mollweide projection).

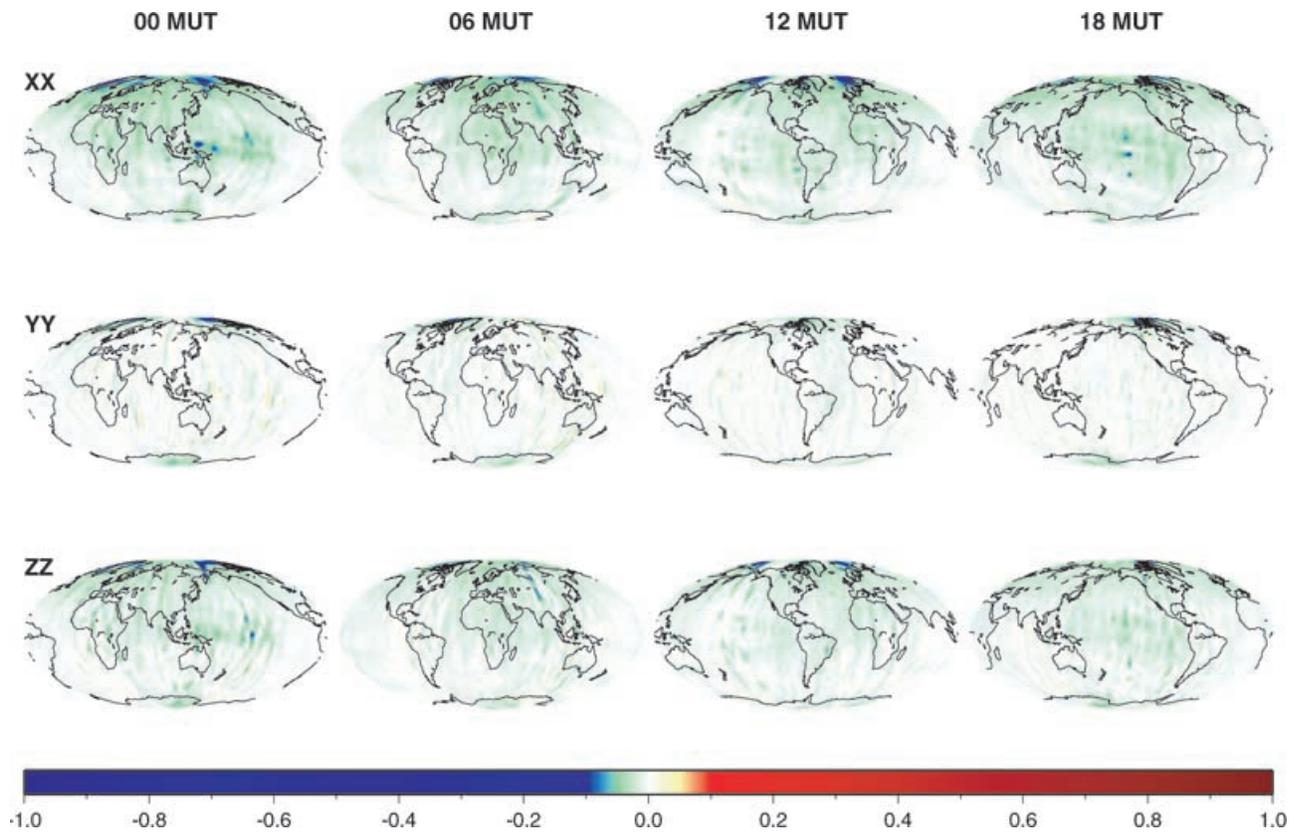


Figure 13. Correlation maps of like vector components from lithospheric and ionospheric fields at  $r = 6771.2$  km on 1980 June 21 centred on noon MLT, but for different MUTs (Mollweide projection).

the weaker structure. Most features appear to be more closely fixed in MLT than in the previous case, and the strongest features are in the northern polar regions of  $XX$  and  $ZZ$ . The more intense localized spots in the Pacific basin of the  $XX$  and  $ZZ$  maps do not appear to correlate with actual features in the lithospheric anomaly maps. However, it must be reiterated that the correlations between the magnetosphere and the ionosphere, and between the lithosphere and the ionosphere are weak at best.

### 5.3 Resolution and calibration

Several quantities related to parameter resolution and error calibration have been computed and are listed in Table 4. They reflect both the data subsets defined in Table 3 and the various norms discussed in Section 4.1. These quantities include two additional classes of misfits: data subset misfits,  $s_{d_i}^2$ , which measure the efficiency of the model in fitting data in that subset; and prior misfits,  $s_{a_i}^2$ , which measure the departure of the model from its  $i$ th *a priori* preferred state per DOF in that norm. They are defined as

$$s_{d_i}^2 = \frac{[\mathbf{d}_i - \mathbf{a}_i(\bar{\mathbf{x}})]^T \mathbf{W}(\mathbf{d}_i - \mathbf{a}_i(\bar{\mathbf{x}}))}{N_i - \text{tr}[\mathbf{R}_{d_i}]}, \quad (62)$$

$$s_{a_i}^2 = \frac{\lambda_i Q_i(\bar{\mathbf{x}})}{M_i - \text{tr}[\mathbf{R}_{a_i}]}, \quad (63)$$

with

$$\mathbf{R}_{d_i} = \mathbf{C}_x \mathbf{A}_i^T \mathbf{W} \mathbf{A}_i, \quad (64)$$

$$\mathbf{R}_{a_i} = \lambda_i \mathbf{C}_x \Lambda_i, \quad (65)$$

where  $N_i$  is the number of observations in the  $i$ th data subset,  $M_i = \dim \mathbf{x}_{a_i}$ , and  $\mathbf{A}_i$  is the submatrix of  $\mathbf{A}$  associated with the

$i$ th data subset. In the table,  $\mathcal{L}_{r_i}(\bar{\mathbf{x}})$  and  $\mathcal{L}_{e_i}(\bar{\mathbf{x}})$  refer to the numerators of eqs (62) and (63), respectively, and  $E[\cdot]$  is the expectation operator. From the fourth column, which gives the trace of the resolution matrix corresponding to either the  $i$ th data subset,  $\mathbf{R}_{d_i}$ , or the norm,  $\mathbf{R}_{a_i}$ , it can be seen that the data are resolving about 59 per cent of the 16 594 total parameters. Of these, about 36, 58 and 6 per cent are resolved by the observatories, Magsat, and POGO, respectively. As for the norms, about 13, 0.3, 10, 9 and 4 per cent of the parameters are resolved by  $Q_{|\bar{b}_r|}$ ,  $Q_{|\nabla_s^2 \bar{b}_r|}$ ,  $Q_{\|\mathbf{J}_{\text{eq}}\|}$ ,  $Q_{\|\nabla_s^2 \mathbf{J}_{\text{eq}, p>0}\|}$ , and  $Q_{\|\Delta \mathbf{B}_{\text{tid}}\|}$ , respectively, while  $Q_{|J_r|}$  dawn and dusk both resolve about 3 per cent of the parameters.

The data subset and prior misfit factors are listed in the last column. These are ratios of the observed to the expected errors of the model, and hence, give an indication of the relative influence of the information associated with each factor. These factors are not independent and the adjustment of one will affect the others in complicated ways (see Seber & Wild 1989) for discussions on the related topic of the iteratively reweighted least-squares (IRLS) method). However, values above (below) unity would suggest a relative decrease (increase) in influence of that information is warranted to achieve calibrated error bounds. The most noteworthy deviations from unity come from the  $Q_{|\bar{b}_r|}$  and  $Q_{\|\Delta \mathbf{B}_{\text{tid}}\|}$  norms, suggesting that their influence be significantly decreased. However, since  $\mathcal{L}_{\|\Delta \mathbf{B}_{\text{tid}}\|}$  accounts for less than 1 per cent of the total variance, its influence is considered minor. As for  $Q_{|\bar{b}_r|}$ , a reduction in its influence would indeed allow for much better fits to the OHMs used in the model, but the resulting fits to the OAMs would be unrealistically poor. Finally, the total data misfit ( $s^2$ ) is 0.87. This suggests that the overall data uncertainties be reduced by about 7 per cent to achieve calibration. Indeed, this is an efficient field model when considering its data misfit in light of its data-to-parameter ratio of 36 compared with that of the GUFM1 model of Jackson *et al.* (2000) (a misfit of 1.16

**Table 4.** CM3 resolution and calibration information.  $\mathcal{L}_{r_i}$  and  $\mathcal{L}_{e_i}$  are the weighted residual and error variances, respectively;  $\mathbf{R}_{d_i}$  and  $\mathbf{R}_{a_i}$  are the data and norm resolution matrices, respectively;  $N_i$  and  $M_i$  are the dimensions of the data and preferred model state vectors, respectively;  $s_{d_i}^2$  and  $s_{a_i}^2$  are the data and prior misfits, respectively; and  $\text{tr}[\cdot]$  and  $E[\cdot]$  are the trace and expectation operators, respectively.

Data subset	$\mathcal{L}_{r_i}(\bar{\mathbf{x}})$	$N_i$	$\text{tr}[\mathbf{R}_{d_i}]$	$E[\mathcal{L}_{r_i}(\bar{\mathbf{x}})]$	$s_{d_i}^2$
OHM-1AM	94 738.0	81 712	924.9	80 787.1	1.17
OHM-MUL					
$ 90^\circ - \theta_d  < 50^\circ$	115 849.2	166 334	1527.6	164 806.4	0.70
$ 90^\circ - \theta_d  \geq 50^\circ$	194 267.7	197 422	1060.5	196 361.5	0.99
Magsat dusk	23 663.1	39 488	2672.4	36 815.6	0.64
Magsat polar dusk	12 251.0	15 973	212.0	15 761.0	0.78
Magsat dawn	27 327.1	44 136	2572.1	41 563.9	0.66
Magsat polar dawn	14 861.5	16 928	191.9	16 736.1	0.89
POGO decimated	17 367.4	22 685	461.2	22 223.8	0.78
POGO pass	5018.6	6 754	133.5	6,620.5	0.76
Subtotal	505 343.6	591 432	9756.2	581 675.8	0.87
Norm	$\mathcal{L}_{e_i}(\bar{\mathbf{x}})$	$M_i$	$\text{tr}[\mathbf{R}_{a_i}]$	$E[\mathcal{L}_{e_i}(\bar{\mathbf{x}})]$	$s_{a_i}^2$
$Q_{ \bar{b}_r }$	25 839.3	2535	2140.6	394.4	65.51
$Q_{ \nabla_s^2 \bar{b}_r }$	6795.1	2535	45.3	2489.7	2.73
$Q_{\ \mathbf{J}_{\text{eq}}\ }$	11 108.1	5520	1591.4	3928.6	2.83
$Q_{\ \nabla_s^2 \mathbf{J}_{\text{eq}, p>0}\ }$	1149.3	4910	1448.9	3461.1	0.33
$Q_{\ \Delta \mathbf{B}_{\text{tid}}\ }$	2907.6	740	612.9	127.1	22.88
$Q_{ J_r }$ dawn	743.8	1044	474.4	569.6	1.31
$Q_{ J_r }$ dusk	744.6	1044	524.2	519.8	1.43
Subtotal	49 288.0	18 328	6837.8	11 490.2	4.29
Grandtotal	554 631.5	609 760	16 594.0	593 166.0	0.94

and a data-to-parameter ratio of 10) and the OIFM model of Olsen *et al.* (2000) (a misfit of 1.01 and a data-to-parameter ratio of 34).

## 6 DISCUSSION

In this section the examination of the CM3 model moves to more physical grounds. The salient features of the model are compared with, and assessed in light of, other works and with the known physics of the near-Earth magnetic field.

### 6.1 Previous comprehensive models

Before a discussion of the various source fields is undertaken, it is instructive to compare the unweighted residual statistics of the CM3 model with its predecessor GSFC(8/95-SqM), introduced in Section 1.3 and for which details are given in Sabaka *et al.* (2000) and in the original paper. The statistics for this model, as well as for the GSFC/CU(12/96) model introduced in Section 4.3, are given in Table 3. Note that only the satellite data sets are identical between CM3 and GSFC/CU(12/96), otherwise the statistics are arranged by qualitatively similar groupings. With the exception of the high-latitude  $Z$  component for CM3, the  $\sigma_r$  for all components of the OHMs generally decrease with increasing model sophistication. The CM3  $Z$  degradation is caused by SV smoothing in  $B_r$ . This is slightly misleading, however, since significant OHM-MUL data discrepancies were detected and corrected between the GSFC/CU(12/96) and CM3 analyses (note the differences in data counts). Since errors in the OAM data were initially suspected, these were replaced by OHM-1AM values. This also expedited the analysis since OAMs are difficult to analyse properly if the integrated nature of their measurements are to be treated properly. The results are that the CM3 fits for OHM-1AM and OHM-MUL are very consistent as well as far superior to the OAM fits of GSFC(8/95-SqM) and GSFC/CU(12/96). The CM3 fits to the OAMs will be discussed shortly. The low-/mid-latitude Magsat data sets also exhibit a decrease in  $\sigma_r$  as the models progress, even when considering the corrected versus uncorrected dusk data sets. For CM3,  $\sigma_r$  for the vector components of the mid-/low-latitude Magsat data are actually less than the 6 nT fluxgate accuracy quoted by Langel & Hinze (1998). However, a more modern treatment, which takes into account such things as attitude anisotropy (see Holme & Bloxham 1996), would show that this is probably an overestimate of the actual error. Therefore, fitting better than 6 nT does not imply an overfitting of the data. The  $\sigma_r$  for the POGO data sets also decrease dramatically from GSFC/CU(12/96) to CM3. This is because the SV influences POGO much more than Magsat and the SV is much improved because of the OHM data error corrections. As expected, the CM3 model is providing superior data fits to those of its predecessors.

### 6.2 Main field secular variation

It is essential that the CM3 SV model be plausible not only because the disjoint Magsat and POGO data envelopes must be correctly linked for proper parameter determination, but because the entire time domain of the model is of scientific interest. To assess the SV, the model will be compared with the GUFM1 model of Jackson *et al.* (2000) in terms of Gauss coefficients (eq. 5) and fits to OAMs, which were not used when deriving the CM3 model. The GUFM1 model, the successor to the UFM1 and UFM2 models of Bloxham & Jackson (1992), spans four centuries (1590–1990) and is derived from a vast historical data compilation. Because the main focus of this model is describing the SV at the CMB, it is well suited as a

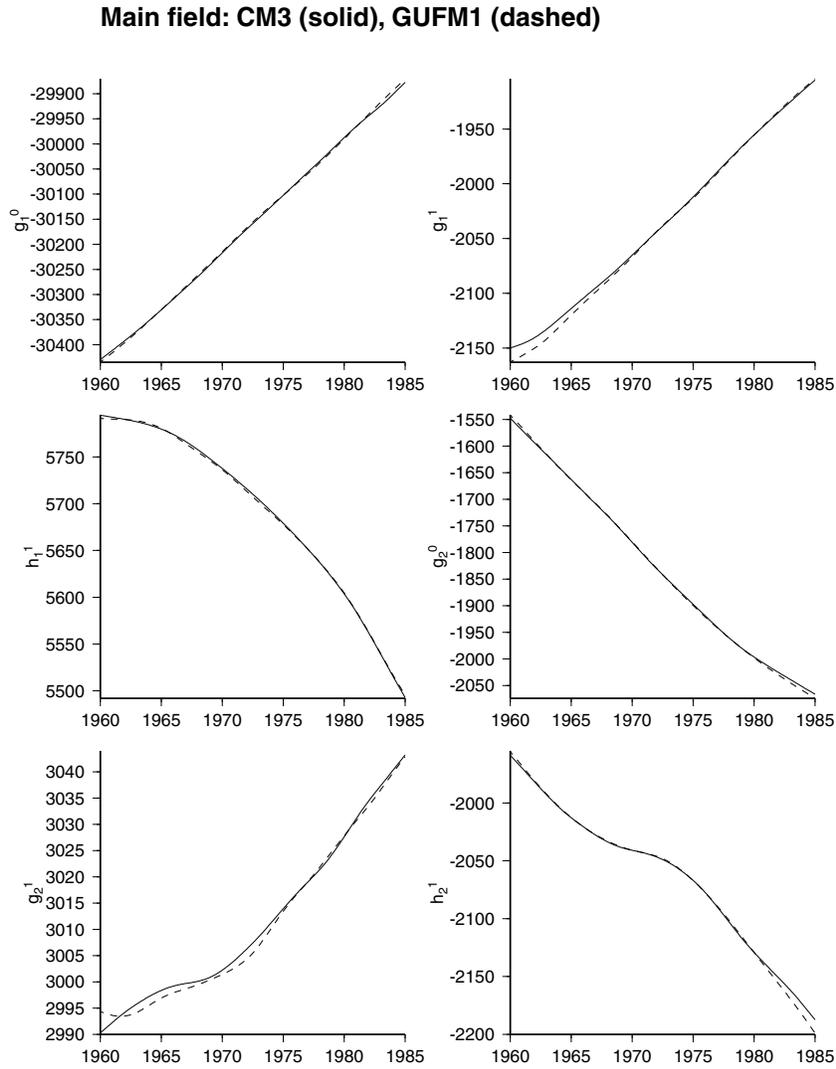
standard for comparison. The model considers internal field sources only and represents the Gauss coefficients as  $B$ -spline expansions in time. In Fig. 14, the first six Gauss coefficients from CM3 (solid) and GUFM1 (dashed) are shown over the time span of CM3. The agreement is quite good, but deviations do occur mostly near the boundaries of the interval. Fig. 15 shows the rates of change of these same coefficients. Both models show the same SV to first order, however, there are some significant differences. Recall that GUFM1 includes no external field contributions nor their induced effects. It is therefore of interest to see whether at least some of the differences may be attributable to aliased long-period induction effects in GUFM1. Of course, this is as much a test for CM3 in its ability to separate core and external time variations. The axial dipole term, in dipole coordinates, from CM3 was subtracted from that of GUFM1 through time and is shown in Fig. 16. It was felt that this component might reveal an aliased signal if one existed. While one could argue from visual inspection that power exists at about the 20 and perhaps even the 10 yr period, a comparison with the predicted signal from the induced portion of CM3 was unconvincing in accounting for the bulk of this difference. Furthermore, this difference is no doubt sensitive to the amount of damping applied to the SV.

The quality of the SV model may also be addressed by its ability to predict long time-series of observatory data, such as the OAMs, which were not analysed in the model. Table 5 compares the rms predictions with 4047 OAM vector observations from GUFM1 and CM3. An additional column shows the effects of not including external field contributions from CM3. The rms values are indeed very close, but the full CM3 model does better for the  $X$  and  $Z$  components. Note that external fields, particularly the magnetospheric field, provide the greatest improvement in the  $X$  component for CM3. Additional measures of the CM3 SV model are provided in Table 6.

### 6.3 Core at 1980 and lithospheric fields

The  $R_n$  spectrum of Lowes (1974) and Mauersberger (1956), which measures the mean-squared magnitude of the internal field over a sphere at a particular epoch per spherical harmonic degree, provides useful information concerning global properties of the spatial character of the model. The spectrum for CM3, computed at  $r = 6371.2$  km and epoch 1980, shows the expected distinct change in slope of  $\ln(R_n)$  around  $n = 14$ , where the core gives way to crustal domination, as noted by Langel & Estes (1982). However, there also appears to be a distinct noise ‘floor’ emerging prior to  $n = 50$ . This type of behaviour is also reported in the spectrum of the M07AV6 model of Cain *et al.* (1989b). This suggests at least three spectral regimes: low-degree core-dominated ( $S_1$ ), mid-degree crustal-dominated ( $S_2$ ) and high-degree noise-dominated ( $S_3$ ). Since the assignment of boundaries between the regimes can be somewhat subjective, especially when done visually, a more objective approach was developed in which a three-segment best-fitting linear piecewise regression (BFLPR) to unweighted  $\ln(R_n)$  (excluding  $\ln(R_1)$ ) was performed as a function of degree partitioning, i.e. the degree boundaries were chosen that minimized the total misfit. Note that a segment from one regime was not extrapolated into the other regimes during the fitting procedure; a process that does not reflect the true spectral overlap of the sources. However, for purposes of comparison this will suffice. The resulting BFLPR is

$$\ln(R_n) = \begin{cases} (-1.27 \pm 0.07)n + (20.8 \pm 0.7) & \text{for } S_1, \\ (0.05 \pm 0.02)n + (1.8 \pm 0.7) & \text{for } S_2, \\ (0.11 \pm 0.03)n - (1.0 \pm 1.7) & \text{for } S_3, \end{cases} \quad (66)$$



**Figure 14.** First six Gauss coefficients from CM3 (solid) and the GUFM1 (dashed) model of Jackson *et al.* (2000), all in nT, over the time span of CM3.

where  $S_1$ ,  $S_2$  and  $S_3$  are found to correspond to degree ranges 2–14, 15–42 and 43–65, respectively. This fit is shown, along with the raw  $R_n$  values, in the top panel of Fig. 17. Points of intersection between regressions for  $S_1$  and  $S_2$ ,  $S_1$  and  $S_3$ , and  $S_2$  and  $S_3$  occur at  $n = 14.4$ , 15.7 and 43.1, respectively, as compared with the  $S_1$  and  $S_2$  value of  $n = 14.2$  given by Cain *et al.* (1989b). Under the assumptions of optimal (Wiener) filtering (Press *et al.* 1992), the noise spectrum is considered uncorrelated with that of the core and the crust and is hypothesized to follow the same trend at lower degrees as established at higher degrees. The bottom panel of Fig. 17 shows the BFLPR to  $\ln(R_n)$ , over the same core- and crustal-dominated regimes of the top panel, after subtraction of  $R_n$  noise values extrapolated from the linear regression for the high-degree spectrum. These regressions are given by

$$\ln(R_n) = \begin{cases} (-1.27 \pm 0.07)n + (20.8 \pm 0.7) & \text{for } S_1, \\ (-0.010 \pm 0.023)n + (2.7 \pm 0.7) & \text{for } S_2. \end{cases} \quad (67)$$

While the noise correction has had a negligible effect on the core-dominated spectrum, it has removed practically all of the slope in the linear regression for the crustal-dominated portion, rendering it nearly level. The radii,  $R'$ , at which the spectra become level may

be an indication of the maximum depth of the current source layers associated with that part of the magnetic field, and for a power law of the form  $R_n = c(d)^n$  at radius  $R$  it can be shown that (Cain *et al.* 1989b)

$$R' = R\sqrt{d}, \quad (68)$$

where  $\ln(d)$  is the slope of the linear regression. Applying this to the corrected core spectrum yields a levelling depth of  $116 \pm 125$  km below the seismic core–mantle boundary at  $r = 3485$  km. This falls generally in between the values of 174 km given by Langel & Estes (1982) and  $80 \pm 46$  km given by Cain *et al.* (1989b), although the unweighted error envelope encompasses these values as well as several kilometres above the CMB. The levelling depth for the crustal spectrum is  $31 \pm 74$  km below the Earth's surface, where the Cain *et al.* (1989b) value is 21 km. This exercise, however, should not be taken out of context given the overly simplifying assumptions that have been made. There is evidence, for instance, from the statistical models of crustal magnetization of Jackson (1994) that the crustal power may be increasing with degree over this range and only at higher degrees ( $n = 200$ ) does it begin to fall-off. In addition, as mentioned earlier, spectral overlap was not considered when establishing the regimes. However, when the overlap is properly accounted for, the

Secular Variation: CM3 (solid), GUFM1 (dashed)

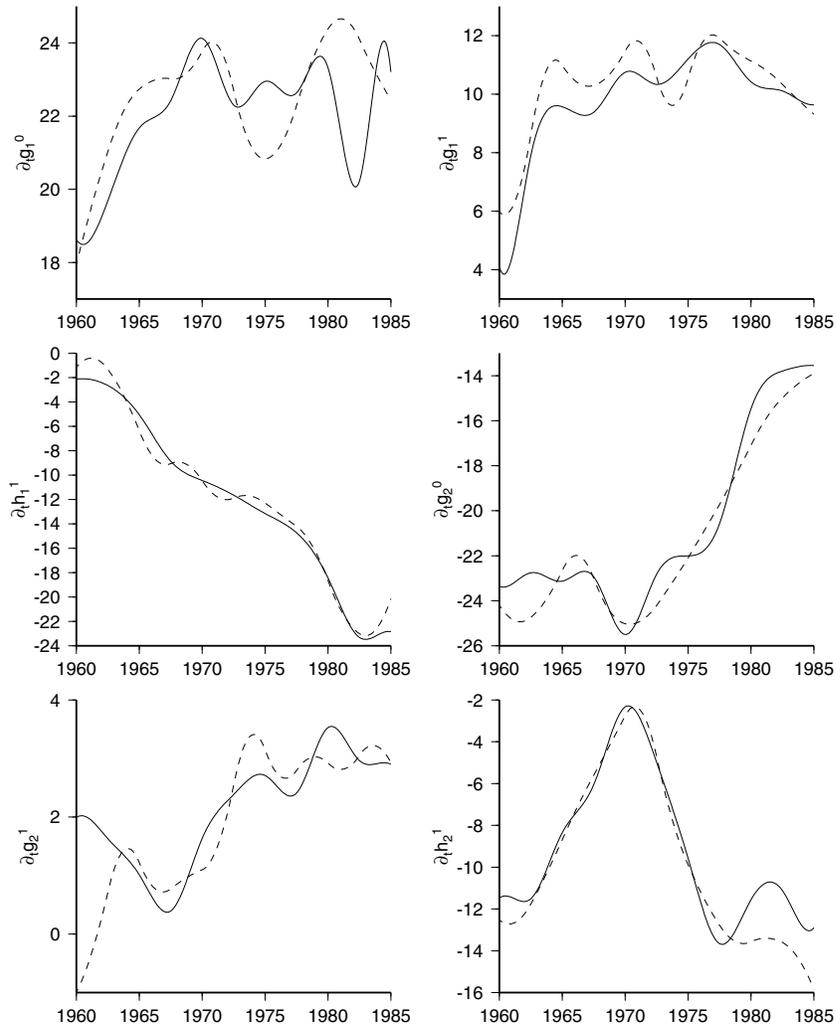


Figure 15. Rate of change of first six Gauss coefficients from CM3 (solid) and the GUFM1 (dashed) model of Jackson *et al.* (2000), all in nT yr<sup>-1</sup>, over the time span of CM3.

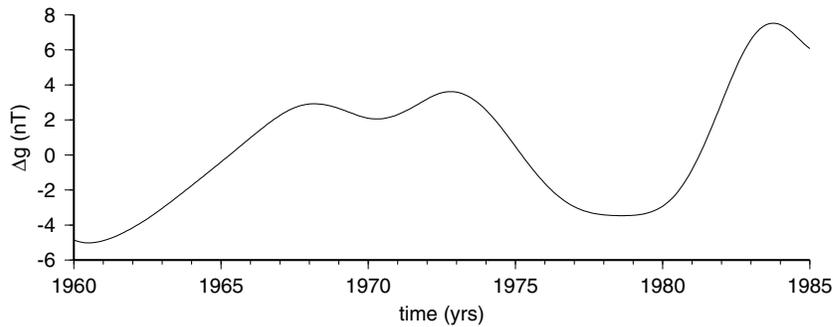


Figure 16. Axial dipole term ( $g_1^0$ ), in dipole coordinates, of CM3 subtracted from that of GUFM1.

levelling depth for the crustal spectrum becomes  $122 \pm 37$  km above the Earth's surface; a value close to the ionospheric *E*-region sheet current height. For comparative purposes, however, these exercises show that the  $R_n$  spectrum of CM3 is in good agreement with previous work, but see Walker & Backus (1997) for a more sophisticated treatment.

In light of the discussion on  $R_n$ , the lithospherically dominated portion of the internal field model is taken to be  $n = 15-42$  and the associated  $B_r$  map at satellite altitude is shown in Fig. 18. The  $\Delta Z$  map for  $n = 15-65$  of Ravat *et al.* (1995) (referred to as the RLPAA map) will be used for comparison. It is derived from Magsat data using various data processing techniques as well as techniques for modelling

**Table 5.** Comparison of rms predictions to OAM data (units in nT).

Component	Number	rms		
		GUFM1	CM3	
			All	No external
<i>X</i>	4047	17.71	17.48	18.09
<i>Y</i>	4047	21.27	21.45	21.47
<i>Z</i>	4047	24.55	24.49	24.53

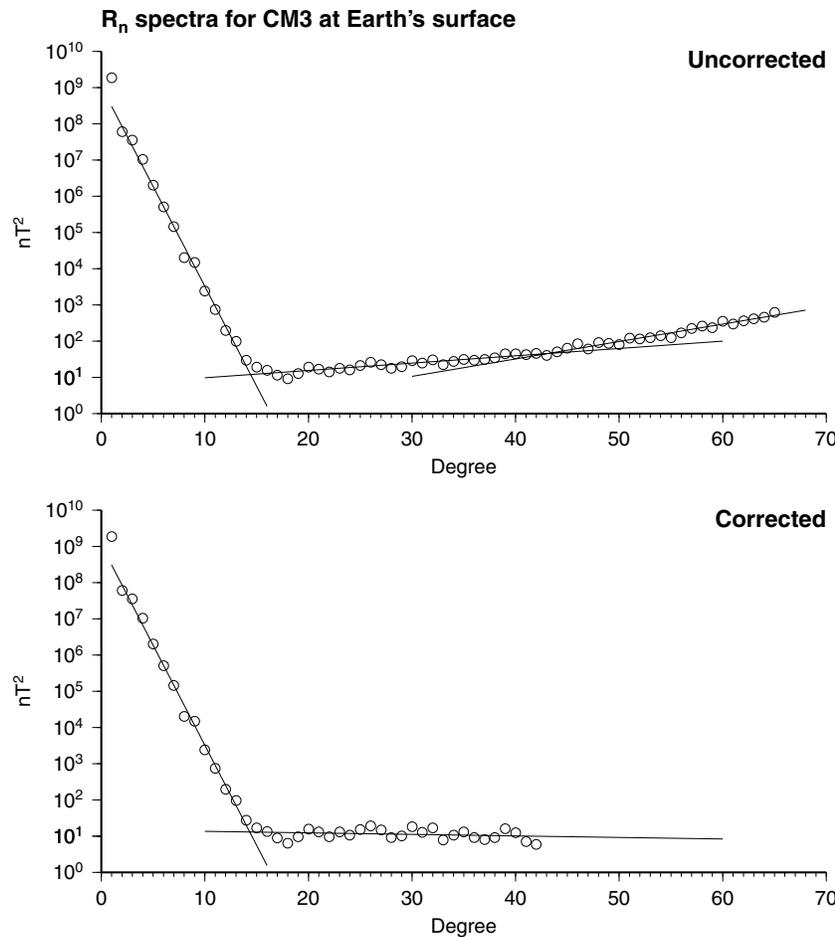
**Table 6.** CM3 secular variation measures at the CMB from 1960 to 1985.

rms radial secular variation	1220	nT yr <sup>-1</sup>
rms radial secular acceleration	28	nT yr <sup>-2</sup>
rms surface Laplacian of radial secular variation	$1.2 \times 10^{-2}$	nT yr <sup>-1</sup> km <sup>-2</sup>

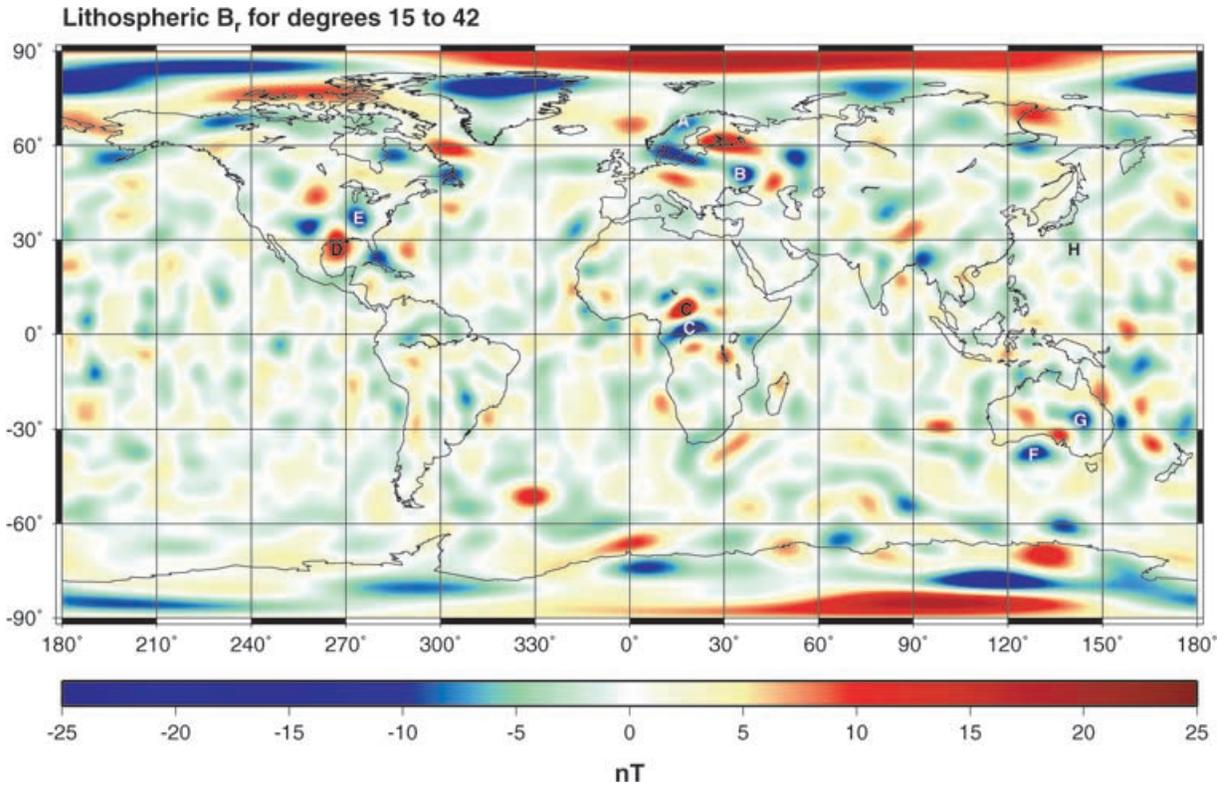
and removing ionospheric fields, and a covariant spherical harmonic analysis procedure to isolate common dawn and dusk lithospheric anomaly features. First, both maps are in excellent agreement with regard to the shapes and locations of the major well-known anomalies such as Kiruna (A), Kursk (B), Bangui (C), Gulf of Mexico (D), Kentucky (E), Australian Bight (F), etc. (letters identify the anomaly locations in Fig. 18). In general, however, the anomaly magnitudes

from CM3 appear larger, as much as 30–50 per cent in some cases. This may be an affect of applying Kaiser filtering (Kaiser 1974) with a cut-off of 12 000 km to the equatorial/mid-latitude Magsat vector data used in deriving RLPAA, resulting in diminished amplitudes. This type of along-track filtering may also remove north–south information from Magsat passes. Some evidence of this may be seen, for instance, when comparing the maps in the eastern Australian basin region (G) (see Purucker *et al.* 1997) and the region of the Izu–Bonin subduction zone (H) (30°N, 140°E). Other CM3 north–south lineations, such as those in the South American mid-continent region just south of the dip equator, may be manifestations of external field contamination, though it is unclear at this point.

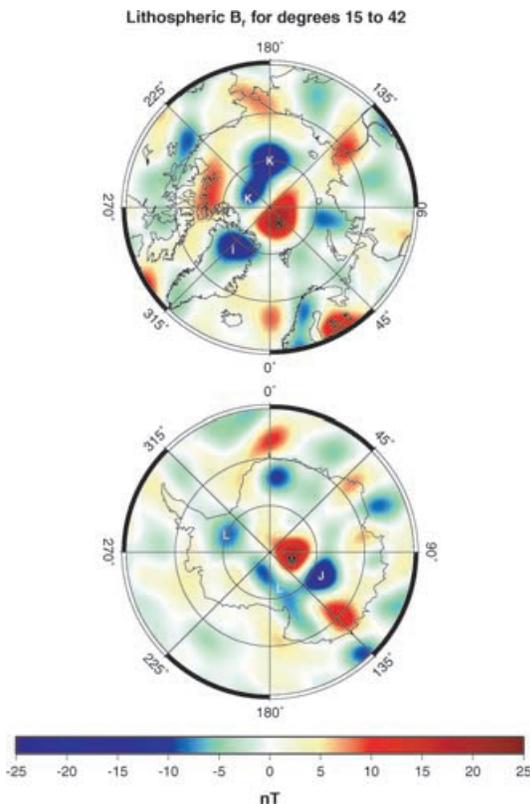
To facilitate the comparison at high latitudes,  $B_r$  polar maps above 60°N and below 60°S are shown in Fig. 19. Because the RLPAA map is derived from Magsat data alone, it is not considered valid above about 83°N or below about 83°S. However, within the region of Magsat coverage, the maps agree very well in general shape and location of the major known anomalies. Again, the anomaly intensities run generally higher for the CM3 maps. There are a couple of notable differences, such as the slight shift of the northern Greenland anomaly (I) from a position on the northwest coast in RLPAA to a more north-central position in CM3, and the emergence of a significant negative anomaly centred over the south magnetic dipole position (J) (79°S, 109°E) in CM3 that is absent in RLPAA. The



**Figure 17.** A comparison of  $R_n$  spectra for the CM3 model at  $r = 6371.2$  km for epoch 1980 corrected (bottom) and uncorrected (top) for high-degree noise contamination. The solid lines in the uncorrected plot are the three best-fitting linear piecewise regressions (BFLPR) to  $\ln(R_n)$  as a function of degree partitioning. The solid lines in the corrected plot are the BFLPR, over the same low- and mid-degree segments, to  $\ln(R_n)$  after subtraction of values extrapolated from the linear regression for the high-degree segment in the uncorrected spectrum.



**Figure 18.** Global map of the CM3 lithospheric contribution (degrees 15–42) to  $B_r$  on the sphere  $r = 6771.2$  km (cylindrical equidistant projection). Letters identify particular anomalies referenced in the discussion.



**Figure 19.** Polar maps above  $60^\circ\text{N}$  (top) and below  $60^\circ\text{S}$  (bottom) of the CM3 lithospheric contribution (degrees 15–42) to  $B_r$  on the sphere  $r = 6771.2$  km (stereographic projections). Letters identify particular anomalies referenced in the discussion.

polar maps also reveal quite a lot of structure in the CM3 lithospheric model that lies near the geographic poles, outside the region of Magsat data coverage. Recall that CM3 includes data from the POGO satellites OGO-2, OGO-4 and OGO-6, for which the orbit inclinations were  $87.3^\circ$ ,  $86.0^\circ$  and  $82.0^\circ$ , respectively. Thus, the gap in polar coverage for CM3 has been reduced to caps of about  $3^\circ$  half-angles. A better standard for comparison in these regions is therefore the POGO-derived anomaly maps of Langel (1990). Since these are maps of scalar anomalies reduced to pole, they should be very close to  $\Delta Z$  and  $\Delta B_r$  near the north and south poles, respectively. The strong positive-negative anomaly pair entwined at the north pole (K) also appears on the POGO map, except here the intensities are stronger than can be accounted for by upward continuation to the 500 km level of the POGO maps. Near the south pole there is a strong negative lineation (L) that parallels the  $135^\circ$ – $315^\circ$  meridian and corresponds very well with the Trans-Antarctic mountain chain. This feature is diffuse and fragmented at best in the POGO map. Again, the strong negative anomaly over the south magnetic dipole (J) is missing in the POGO map. Finally, there is a strong new feature (M) in the CM3 map located at about  $86^\circ\text{S}$ ,  $90^\circ\text{E}$  with a peak value of almost 19 nT. While it has no counterpart in the POGO map, the CM3 model does include POGO measurements over a portion of this region. However, since the POGO data gap also transects this feature, an interpretation must be suspended until further analysis is performed.

#### 6.4 Ionospheric field

The most dominant of the expansion coefficients describing the ionospheric  $E$ -region currents are  $\Re\{\tilde{\epsilon}_{1,0,0}^0\}$ ,  $\Re\{\tilde{\epsilon}_{2,0,1}^1\}$  and  $\Re\{\tilde{\epsilon}_{3,0,2}^2\}$ . At sunspot minimum, where typical values of  $F_{10.7} = 66.0 \times 10^{-22}$   $\text{W m}^{-2} \text{Hz}^{-1}$ , these have values of  $\Re\{\tilde{\epsilon}_{1,0,0}^0\} = -4.0$  nT,  $\Re\{\tilde{\epsilon}_{2,0,1}^1\} =$

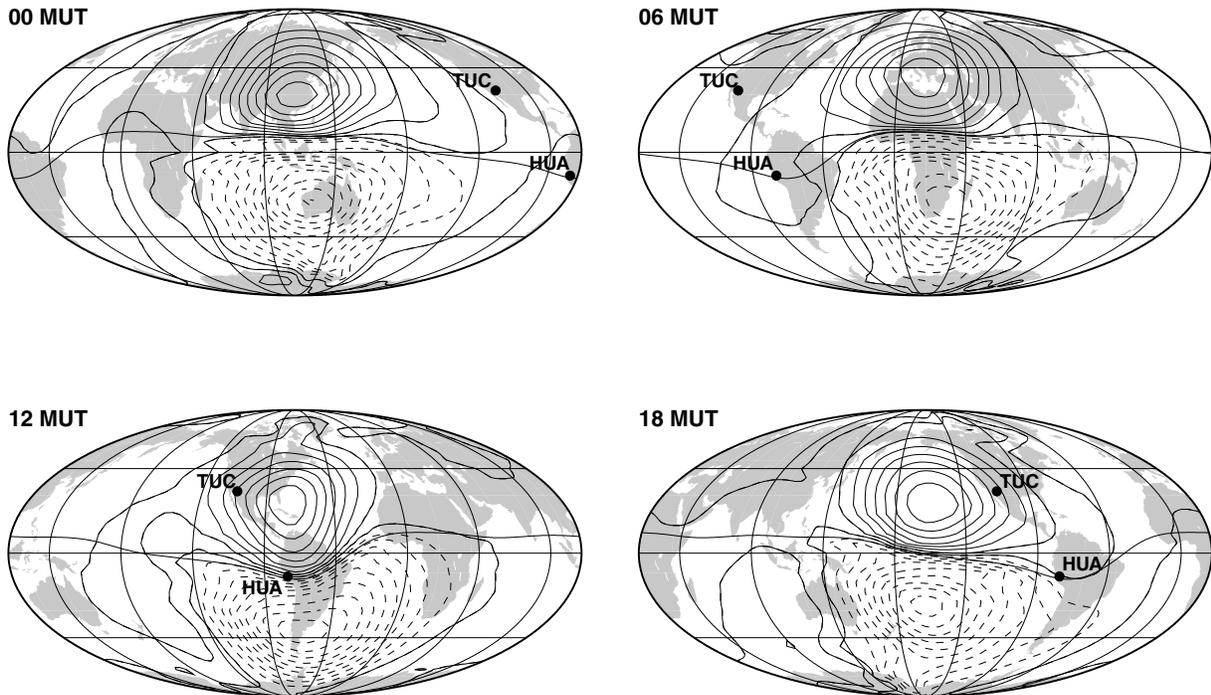
5.6 nT and  $\Re\{\tilde{\epsilon}_{3,0,2}^2\} = -3.2$  nT, while at sunspot maximum the coefficient magnitudes are about twice as large. All are local time terms ( $l = p$ ) with no seasonal variation ( $s = 0$ ), the maximum amplitudes of which are centred at the subsolar point (real parts) and contribute to the two major Sq foci of antipolarity in the northern and southern hemispheres ( $k - l = 1$ ). The first coefficient describes the general polarity between the north and south foci at all local times ( $p = 0$ ), the second describes the regular diurnal variation ( $p = 1$ ) of such a field through the course of a solar day, while the third imparts some semi-diurnal variation ( $p = 2$ ) to this.

The dominance of  $\Re\{\tilde{\epsilon}_{2,0,1}^1\}$  can be seen in global maps of the  $E$ -region equivalent current function  $\Psi$  of eq. (54). In Fig. 20,  $\Psi$  is shown at MLT noon on 1980 March 21, but for different values of MUT. A value of  $F_{10.7} = 140.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$ , an average over the time span of this model, was used to generate the maps. The dual Sq foci are indeed the major features, showing a slightly asymmetric current load (rather than the symmetric one expected for equinox) flowing in oppositely directed vortices in the northern (counterclockwise) and southern (clockwise) hemispheres in accordance with eq. (53). The total current flowing in the northern (southern) vortex is the same (slightly lower) than that reported by Malin & Gupta (1977), who also included  $p = 0$  terms. However, their data were acquired during IGY when  $F_{10.7}$  levels were elevated and their analysis was unable to distinguish between ionospheric and magnetospheric effects. Besides these, there are several other items of interest: a significant decrease in  $\|\nabla\Psi\|$  exists for much of the darkside hemisphere at all MUT, indicating that the  $Q_{\parallel J_{\text{eq}}}$  constraint is effective; the boundary between the two foci is coincident with the dip equator at all MUT, i.e. current flows tangential to the dip equator at and near local noon, thus affirming the utility of the QD constraints; there is also some amplitude and shape modulation with

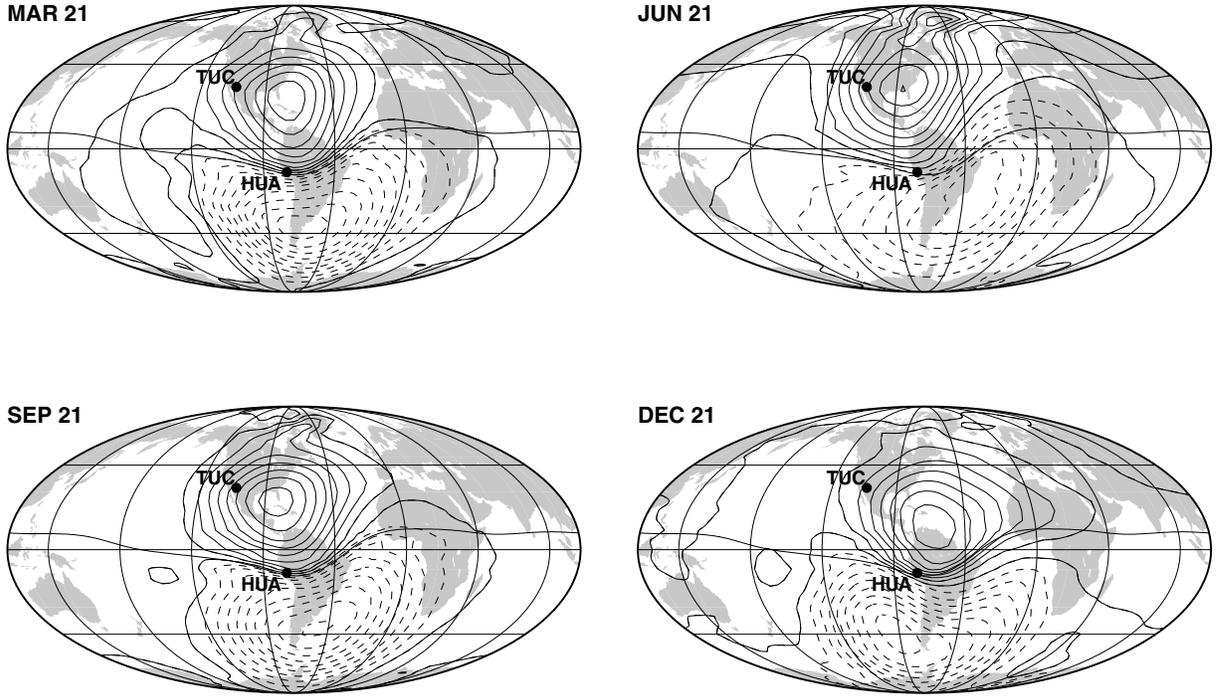
MUT that is beyond what is inherent in the QD constraints and which is attributable to non-local time variation; and there is a marked increase in  $\|\nabla\Psi\|$  parallel to the dip equator at and near local noon for all MUT, revealing an enhanced eastward current flow, which is in fact the EEJ.

In Fig. 21,  $\Psi$  is shown at noon MLT in the centre of the figure, but for the equinoxes and solstices. A distinct seasonal variation is evident, as expected. Very clearly, the summer focus occurs at an earlier local time than the winter focus—a known feature that persists at all MUTs. Though not apparent in Fig. 21 at noon MUT, similar plots at midnight MUT show a minimum northern (southern) focal current intensity in December (June) and a maximum northern focal intensity in June, as expected, but a maximum southern focal intensity in September. This general peculiarity in focal intensity variation with season is probably attributable to the  $Q_{\parallel J_{\text{eq}}}$  constraint that imposes circuit closure on the dayside hemisphere and, together with the static ionospheric field, may very well be modifying the expected variation. Seasonal oscillations in  $\Psi$  focal intensities with respect to a preliminary CM3 model including no  $Q_{\parallel J_{\text{eq}}}$  constraint are indeed of higher amplitude, confirming this idea.

These  $E$ -region current signatures are also clearly seen in the field intensity of the POGO pass in Fig. 8 and the field components of Magsat dusk pass 263 in Figs 8 and 9. Note the accentuation of the  $B$  minimum at noon over the dip equator as well as for  $X$  from the dusk extent of the EEJ. The OHM component plots from Tucson (Fig. 5) and Huancayo (Fig. 6) also reveal something of the daily, seasonal and spatial structure of the  $E$ -region currents. Noting that the daily segments begin at 00 UT, it can be seen that Tucson  $X$  and  $Z$  both have a single negative spike at local noon, while  $Y$  spikes positive just before and negative just after local noon. This agrees with the location of Tucson being below and just north of the northern focus



**Figure 20.** Global maps of ionospheric  $E$ -region equivalent current function  $\Psi$  for 1980 March 21 (Mollweide projection). A value of  $F_{10.7} = 140.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$ , an average over the time span of this model, was used to generate the maps. Recall from eq. (54) that  $\Psi$  is defined on the sphere  $r = 6481.2$  km. Each of the four panels is centred on noon MLT, but for different MUTs. The associated induced contribution is not included here. A 20 kA current flows between the contours. Locations of the Tucson (TUC) and Huancayo (HUA) observatories are shown.



**Figure 21.** Global maps of ionospheric  $E$ -region equivalent current function  $\Psi$  defined on the sphere  $r = 6481.2$  km (Mollweide projection). A value of  $F_{10.7} = 140.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$ , an average over the time span of this model, was used to generate the maps. Each of the four panels is centred on noon MLT and MUT, but for different seasons, i.e. 1979 December 21 and March 21, June 21 and 1980 September 21. The associated induced contribution is not included here. A 20 kA current flows between the contours. Locations of the Tucson (TUC) and Huancayo (HUA) observatories are shown.

of  $\Psi$  in Fig. 20. At Huancayo,  $X$  spikes strongly positive (the EEJ) and  $Z$  spikes negative at local noon, which agrees with its location below and perhaps slightly north of the dip equator as shown in Fig. 20. In addition, Tucson  $Y$  and  $Z$  and all Huancayo components indicate an intensification of the ionospheric contribution to the diurnal signal during their respective summers.

### 6.5 Magnetospheric field

As expected, the static term along the dipole axis,  $\Re\{\mu_{1,0,0}^0\}$ , is by far the most dominant, having a value of 21.4 nT. The magnitudes of the annual and semi-annual variations along this axis are about 13 and 7 per cent of this magnitude, respectively. Although the mean tilt of the non- $D_{st}$ -dependent magnetospheric dipole field over season is only about  $2.7^\circ$  towards MLT noon, i.e.  $\Re\{\mu_{1,0,1}^1\} = 1.0$  nT, the seasonal fluctuation about this mean is much larger, being predominantly in the annual term, with the peak tilt magnitude being  $15.3^\circ$  at the solstices, i.e.  $|\Re\{\mu_{1,1,1}^1\} + \Re\{\mu_{1,-1,1}^1\}| = 4.8$  nT.

As for higher multipole terms, some have been detected in previous studies. Malin & Mete Isikara (1976) studied the annual variation of midnight values of observatory data and found contributions to  $\mu_{1,1,0,g}^0$ ,  $\mu_{2,1,0,g}^0$  and  $\mu_{2,1,1,g}^1$  (where the subscripted ‘g’ indicates coefficients in geographic spherical-polar coordinates) that they attributed to seasonal movement of the ring-current relative to the equatorial plane. Note that their analysis is unable to distinguish annual variation of the magnetospheric part ( $\mu_{nsp,g}^m$ ) from that of the ionospheric part ( $\epsilon_{nsp,g}^m$ ). If the real part of a particular coefficient, say  $\Re\{c_{np}^m(t)\}$ , is Fourier analysed for annual periodicity such that

$$\Re\{c_{np}^m(t)\} = A \cos(\omega_s t - \varphi), \quad (69)$$

then Table 7 summarizes their results under the M&MI column. Comparable results are also shown for CM3 and a preliminary CM3-

type model without  $Q_{\parallel \Delta \mathbf{B}_{\text{tid}}}$  smoothing, denoted CNXS; both for magnetospheric terms alone and for the sum of magnetospheric and ionospheric terms of similar indices. A value of  $F_{10.7} = 140.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$  was used to compute the ionospheric coefficients, and all coefficients are in geographic spherical-polar coordinates. Comparing coefficients of the form  $\Re\{\mu_{np,g}^m + \epsilon_{np,g}^m\}$ , it can be seen that the phases,  $\varphi$ , are in fair agreement (within about  $40^\circ$ ) with very tight agreement for  $\Re\{\mu_{2,0,g}^0 + \epsilon_{2,0,g}^0\}$  and less so for  $\Re\{\mu_{1,0,g}^0 + \epsilon_{1,0,g}^0\}$ . The amplitudes,  $A$ , of CNXS are about half those of M&MI, while those of CM3 are even less. This decreased amplitude in the annual variation of the CM3-type models is probably due in part to the  $Q_{\parallel \mathbf{J}_{\text{eq}}}$  smoothing that operates on ionospheric seasonal variations at all  $p$  values. However, since the magnetospheric annual variations are not compensating for this smoothing, then perhaps the bulk of the M&MI amplitudes is actually of ionospheric origin. The apparent degradation from CNXS to CM3 with respect to the M&MI findings is probably the result of the  $Q_{\parallel \Delta \mathbf{B}_{\text{tid}}}$  smoothing, which is damping

**Table 7.** Magnetospheric annual variation (nT) from Malin & Mete Isikara (1976) compared with CM3-type models. A value of  $F_{10.7} = 140.0 \times 10^{-22} \text{ W m}^{-2} \text{ Hz}^{-1}$  was used to compute the ionospheric coefficients, and all coefficients are in geographic spherical-polar coordinates.

$\Re\{c_{np}^m(t)\}$	M&MI		CM3		CNXS	
	$A$	$\varphi$	$A$	$\varphi$	$A$	$\varphi$
$\Re\{\mu_{1,0,g}^0 + \epsilon_{1,0,g}^0\}$	3.8	$190^\circ$	1.6	$231^\circ$	1.9	$220^\circ$
$\Re\{\mu_{2,0,g}^0 + \epsilon_{2,0,g}^0\}$	6.0	$352^\circ$	2.8	$349^\circ$	3.6	$350^\circ$
$\Re\{\mu_{2,1,g}^1 + \epsilon_{2,1,g}^1\}$	1.3	$20^\circ$	0.2	$18^\circ$	0.4	$31^\circ$
$\Re\{\mu_{1,0,g}^0\}$	–	–	2.7	$222^\circ$	3.1	$220^\circ$
$\Re\{\mu_{2,0,g}^0\}$	–	–	0.1	$353^\circ$	1.4	$343^\circ$
$\Re\{\mu_{2,1,g}^1\}$	–	–	0.04	$81^\circ$	0.4	$59^\circ$

**Table 8.** Magnetospheric daily variation (nT) from T87We Peredo *et al.* (1993) compared with CM3-type models.

	$\Re\{\mu_n^m(t_s, t_{\text{mut}})\}$	Season	T87We ( $K_p = 0$ )		T87We ( $K_p = 3$ )		CM3		CNXS	
			$A$	$\varphi$	$A$	$\varphi$	$A$	$\varphi$	$A$	$\varphi$
Diurnal	$\Re\{\mu_1^1\}$	December	3.0	180°	3.3	180°	4.0	185°	3.7	182°
		Equinox	0	–	0	–	0.7	55°	2.1	162°
	$\Re\{\mu_2^1\}$	December	1.2	359°	2.3	359°	0.2	343°	1.3	356°
		Equinox	1.4	357°	2.7	357°	0.3	348°	3.2	0°
	$\Re\{\mu_1^0\}$	December	1.3	181°	1.9	181°	0.1	329°	0.6	218°
		Equinox	0	–	0	–	0.2	97°	0.5	58°
Semi-diurnal	$\Re\{\mu_2^2\}$	December	0.04	359°	0.1	359°	0.1	346°	0.3	328°
		Equinox	0	–	0	–	0.1	338°	2.6	357°
	$\Re\{\mu_1^1\}$	December	0.8	181°	1.1	181°	0.1	171°	0.7	177°
		Equinox	0.7	174°	0.7	174°	0.1	230°	0.6	151°

much of the amplitude of the  $n = 2$  terms in the CM3 model. This is clearly seen by comparing coefficients of the form  $\Re\{\mu_{np,g}^m\}$ ; the amplitudes of  $\Re\{\mu_{2,0,g}^0\}$  and  $\Re\{\mu_{2,1,g}^1\}$  are an order of magnitude less in CM3. If M&MI are correct, then perhaps this damping is excessive, which agrees with the conclusions of Section 5.3.

Olsen (1996) has investigated possible magnetospheric field contributions to daily magnetic field variations. He examined several semi-empirical models of the magnetospheric field, reporting primarily on results from Tsyganenko (1987, 1989) and, in particular, the model designated T87We in Peredo *et al.* (1993). If the real part of a particular coefficient, say  $\Re\{\mu_n^m(t_s, t_{\text{mut}})\}$ , is Fourier analysed for diurnal ( $p = 1$ ) or semi-diurnal ( $p = 2$ ) periodicities in MUT at a fixed season (specified by  $t_s$ ) such that

$$\Re\{\mu_n^m(t_s, t_{\text{mut}})\} = A \cos(\omega_p p t_{\text{mut}} - \varphi), \quad (70)$$

then Table 8 gives a comparison of the amplitude,  $A$ , and phase,  $\varphi$ , of some coefficients of interest from T87We, CM3 and CNXS. The agreement of the amplitudes and phases of  $\Re\{\mu_1^1\}$  and  $\Re\{\mu_2^1\}$  in December and at equinox between the T87We and CNXS models is considered good, except in the case of the diurnal variation of the former coefficient at equinox, which is negligible in the T87We model. The amplitudes and phases deviate significantly more for  $\Re\{\mu_1^0\}$  and  $\Re\{\mu_2^2\}$  during both seasons between these two models; again, the diurnal variation of the former and the semi-diurnal variation of the latter are near zero at equinox in T87We. As for CM3, it agrees well with T87We in its diurnal variation of  $\Re\{\mu_1^1\}$  in December and its diminished amplitude at equinox. However, the level of agreement is much less in all other cases where T87We variations are not zero, particularly the amplitudes. The fact that the agreement is best for the single local time tilted dipole coefficient of the list indicates that the  $\mathcal{Q}_{\parallel \Delta \mathbf{B}_{\text{tid}}}$  smoothing may again be excessive. Hence, the better agreement of the diurnal variation of  $\Re\{\mu_1^0\}$  and the semi-diurnal variation of  $\Re\{\mu_2^2\}$  at equinox for CM3 compared with the zero values for T87We is considered fortuitous.

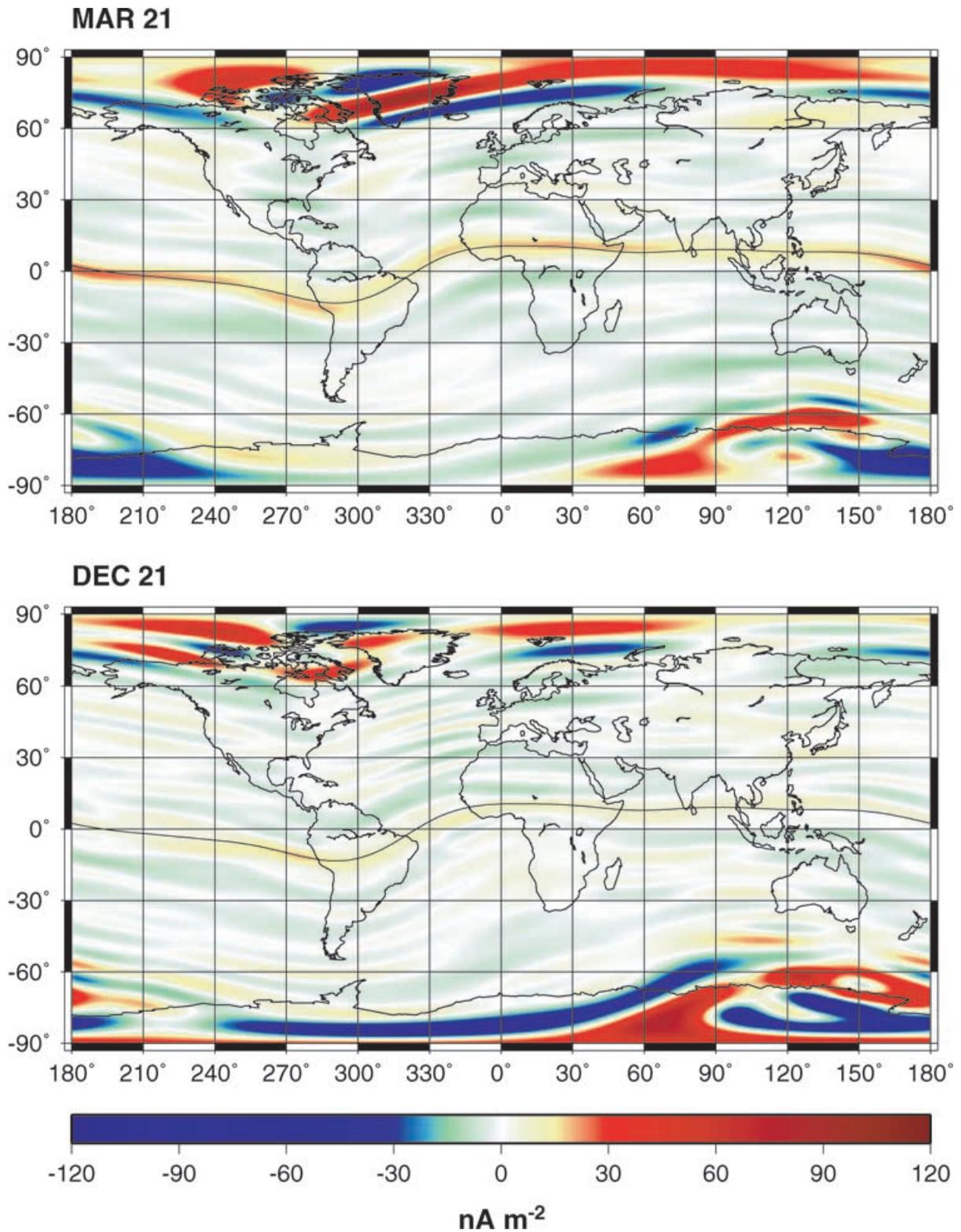
The magnetospheric model predictions of the scalar intensity along the POGO pass shown in Fig. 8 and the vector components along the Magsat dusk pass 263 in Figs 8 and 9 are clearly those of an axial dipole whose moment is aligned, on average, nearly perpendicular to the orbit normals of the satellites. Furthermore, the behaviour of the model as a function of the  $D_{\text{st}}$  index is illustrated in the  $X$  and  $Z$  components of Tucson in 1967 (see Fig. 5) where very large discontinuities are found in the magnetic records between the quietest days chosen for May and June. They are evidently the result of a ring-current adjustment, which is described quite well by  $D_{\text{st}}$ , and hence, the model fits them with little problem.

## 6.6 Fields from ionospheric coupling currents

The radial current density ( $J_r$ ) of the  $F$ -region coupling currents from CM3 at dusk local time is shown in Fig. 12 for March and December 21. Note that this figure does not show  $J_r$  at any particular UT as a function of longitude, but rather, it shows the  $J_r$  predicted to flow at different longitudes at dusk local time. As expected, the largest radial current flows at polar latitudes. During the morning (not shown), currents flow into the ionosphere ( $J_r < 0$ ) at the poleward boundary of the polar oval (region 1 currents), and out of the ionosphere at the equatorial boundary (region 2 currents). During the evening the current direction is reversed. The evening data also show upward currents at the dip equator and downward currents at nearby low latitudes. This is the radial component of the meridional current system of the EEJ, first observed in Magsat data by Maeda *et al.* (1982). There is no evidence for such a current system in the morning sector. A strong seasonal variation can be seen in  $J_r$  between March and December. During the southern summer it is much more intense in the southern polar oval region than in the northern one, while during equinox the intensities are more equal. In addition, the radial component of the meridional current system is better defined during equinox than during solstice. Although these findings are in keeping with those of Olsen (1997a), he finds current intensities that are in general significantly higher than those predicted by CM3. This discrepancy, however, may be attributed to the difference in data selection: all days versus quiet days.

Olsen (1997a) has also detected larger-scale upward currents in the north and downward currents in the south in the evening, and opposite this in the morning, for the December Magsat data. These are obscure in Fig. 22, if present at all, and may probably be a result of the application of the  $\mathcal{Q}_{\parallel J_r}$  smoother at both the dawn and dusk local times, and again to data selection. Such interhemispheric coupling currents are small or absent during the equinoxes and are expected to reverse during the northern summer.

The toroidal field signature of the meridional current system may be clearly seen in the  $Y$  component of Magsat dusk pass 263 in Fig. 9. Because this pass occurs on 1979 November 19, near the southern summer, the current density vortex just to the south of the dip equator is stronger than that just to the north, resulting in a stronger eastward and weaker westward field just south and north of the dip equator, respectively. The CM3 model, in turn, predicts this asymmetry very well. The high-amplitude excursions in the  $X$  and particularly the  $Y$  components at high latitudes are probably also of the field-aligned variety. However, these are also probably very transient in nature, and are thus fitted very poorly by a model of this type, which sees mean temporal and spatial field effects best.



**Figure 22.** Global maps of the radial component of the ionospheric coupling currents ( $J_r$  in eq. 59) at dusk on the sphere  $r = 6821.2$  km for March 21 and December 21 (cylindrical equidistant projection).

## 7 CONCLUSIONS

The paradigm of comprehensive modelling is quite worthwhile, yet quite formidable. Progress has been slow and incremental, and with new satellite missions and new technology on the horizon, a ‘final model’ is simply an ideal to work towards. In this final section an attempt will be made to gauge the position of CM3 and its methodology in the continuum of comprehensive modelling.

### 7.1 New features

Perhaps the best way to assess the progress made in the current phase is to simply compare it with the previous phase. Hence, features of the CM3 model which are new with respect to the GSFC(8/95-SqM) model (Langel *et al.* 1996) will be enumerated in this section.

The main field and attendant SV portions (triple summation in eq. 5) of the models are identical, but CM3 includes a static representation of the high-degree lithospheric field (double summation in eq. 5), which has successfully captured most of the known crustal anomaly fields seen at satellite altitude. In addition, the SV is now much more plausible because of the regularization imparted by the  $Q_{|\hat{b}_{r1}|}$  and  $Q_{|\nabla_{\hat{b}_{r1}}^2 \hat{b}_{r1}|}$  norms.

As for the ionosphere, the CM3 parametrization has higher latitudinal resolution such as to fit the field of the EEJ. Both non-local time terms and QD constraints allow the CM3 field to better respond to the ambient field, especially in terms of the conductivity distribution. Furthermore, the  $Q_{|J_{eq}|}$  constraint injects known physical limits on the conductivity patterns at night time. The GSFC(8/95-SqM) model includes no such features. Rigid contraction and expansion of the ionospheric field in response to solar activity is also found in the CM3 model, but not GSFC(8/95-SqM).

The magnetospheric field of CM3 differs in two major ways from that of GSFC(8/95-SqM): first, it has the possibility of displaying smaller-scale features in both latitude and longitude; and secondly, it contains non-local time terms. Unfortunately, because of separability problems between the magnetospheric and ionospheric fields inherent in the data sets used, the full impact of these differences could not be explored.

Induced fields from time-varying external fields are included in both models. For GSFC(8/95-SqM), the field parametrization is explicit, that is, independent of that of the primary sources. For CM3, the induced field parametrization is coupled with that of the primary sources through an *a priori* conductivity model. The reason for this is to reduce the size of the already large parameter set. Though the independent approach is inevitably of more interest, the coupled approach does allow for very complicated conductivity distributions to be included in the model, via the  $\mathbf{Q}$  matrix of eq. (12), with no additional parameters or computational cost.

The parametrization of fields from ionospheric coupling currents is completely absent in the GSFC(8/95-SqM) model. This is obviously a very significant part of the measured field for satellites moving through these current regions, as can be attested to by  $Y$ -component plots of Magsat dusk data (e.g. Fig. 9). In addition, the CM3 representation is fairly sophisticated in that it includes seasonal variation and a mechanism for conforming to the existing conductivity patterns through the use of QD constraints.

### 7.2 Future work

While the previous section discussed in some sense how far the comprehensive modelling effort has come, this section will hope-

fully give some clue as to how far it has to go, at least in the near future. Much of the future work will focus on data issues: the inclusion of vector data from satellite missions such as Ørsted, CHAMP and others are expected to greatly enhance the validity of the models in local times other than dawn and dusk. Furthermore, these data will aid in the separation of ionospheric and magnetospheric fields. One could also expand the scope of future models by considering measurements from more magnetically disturbed times.

As for parametrization issues: the extension of the main field SV domain to the Ørsted/CHAMP epoch, the inclusion of a  $J_{\theta}$  component to the coupling current density, and a continuous diurnally varying toroidal field will be pursued. More long-term goals include the proper treatment of fields from the ionosphere and associated coupling currents with respect to polar currents that are dependent upon the interplanetary magnetic field, the extension to more magnetically disturbed periods, and the incorporation of more complicated *a priori* conductivity models.

### 7.3 Possible uses

The success of comprehensive modelling is in part driven by its utility to the scientific community. The method of coestimating fields from several sources and its affect on model consistency is of scientific interest in its own right; however, additional merit of the comprehensive models lies in their use as application tools, or reference models. Indeed, with the possible exception of the high-degree lithospheric field where new, physically meaningful features might be found, most source fields are parametrized so as to model the well-known, regular, quiet-time features. Hence, comprehensive models are well qualified for removing known fields from the data so as to not obfuscate that which is unknown. However, one must be cautious when applying them outside their scope, that is, extrapolating them to regimes that were not sampled by the data sets used in deriving them. Uncritical application of CM3 during magnetically disturbed conditions, for instance, is dangerous.

### 7.4 Availability

In accordance with the previous section, a forward modelling code is available, which predicts the various CM3 source fields given spatial and temporal positions and  $D_{st}$  and  $F_{10.7}$  values. This code is in the form of an ANSI standard Fortran subroutine called CM3FIELD. It returns the local (north, east, down) components of  $\mathbf{B}$  in nT on a sphere or on the IAU1966 ellipsoid from internal, primary and induced ionospheric, primary and induced magnetospheric, and dawn or dusk coupling current field sources. Two evaluations of the internal field are accommodated per two given spherical harmonic degree ranges. This is helpful when separate predictions are desired from the core-dominated and lithosphere-dominated portions of the expansion. The CM3 model information is input to the CM3FIELD subroutine on the initial call from a sequentially accessed ASCII file. Both the forward code file and the CM3 model information file are available from the authors by request.

## ACKNOWLEDGMENTS

We would like to thank Art Richmond for his expertise in several external-field-related issues and for help with the initial stages of the manuscript. We also thank Richard Holme and Gauthier Hulot for their thorough reviews. All figures were produced with GMT

(Wessel & Smith 1991). TJS was supported under NASA Contract NAS5-31760.

## REFERENCES

- Arkani-Hamed, J. & Strangway, D.W., 1985a. Intermediate-scale magnetic anomalies of the Earth, *Geophysics*, **50**, 2817–2830.
- Arkani-Hamed, J. & Strangway, D.W., 1985b. Lateral variations of apparent susceptibility of lithosphere deduced from Magsat data, *J. geophys. Res.*, **90**, 2655–2664.
- Arkani-Hamed, J. & Strangway, D.W., 1986. Band-limited global scalar magnetic anomaly map of the Earth derived from Magsat data, *J. geophys. Res.*, **91**, 8193–8203.
- Arkani-Hamed, J., Langel, R.A. & Purucker, M.E., 1994. Magnetic anomaly maps of Earth derived from POGO and Magsat data, *J. geophys. Res.*, **99**, 24 075–24 090.
- Backus, G.E., 1986. Poloidal and toroidal fields in geomagnetic field modeling, *Rev. Geophys.*, **24**, 75–109.
- Barton, C.E., 1997. International geomagnetic reference field: the seventh generation, *J. Geomag. Geoelectr.*, **49**, 123–148.
- Bloxham, J. & Jackson, A., 1992. Time-dependent mapping of the magnetic field at the core–mantle boundary, *J. geophys. Res.*, **97**, 19 537–19 563.
- Bloxham, J., Gubbins, D. & Jackson, A., 1989. Geomagnetic secular variation, *Phil. Trans. R. Soc. Lond., A.*, **329**, 415–502.
- Cain, J.C., Schmitz, D.R. & Muth, L., 1984. Small-scale features in the Earth's magnetic field observed by Magsat, *J. geophys. Res.*, **89**, 1070–1076.
- Cain, J.C., Wang, Z., Kluth, C. & Schmitz, D.R., 1989a. Derivation of a geomagnetic model to  $n = 63$ , *Geophys. J.*, **97**, 431–441.
- Cain, J.C., Wang, Z., Schmitz, D.R. & Meyer, J., 1989b. The geomagnetic spectrum for 1980 and core–crustal separation, *Geophys. J.*, **97**, 443–447.
- Cain, J.C., Holter, B. & Sandee, D., 1990. Numerical experiments in geomagnetic modeling, *J. Geomag. Geoelectr.*, **42**, 973–987.
- Campbell, W.H., 1989. The regular geomagnetic field variations during quiet solar conditions, in *Geomagnetism*, Vol. 3, pp. 385–460, ed. Jacobs, J.A., Academic Press, London.
- Chapman, S. & Bartels, J., 1940. *Geomagnetism*, Clarendon Press, Oxford.
- Cohen, Y. & Achache, J., 1990. New global vector magnetic anomaly maps derived from Magsat data, *J. geophys. Res.*, **95**, 10 783–10 800.
- Counil, J., Cohen, Y. & Achache, J., 1991. The global continent–ocean magnetization contrast, *Earth planet. Sci. Lett.*, **103**, 354–364.
- Hamoudi, M., Cohen, Y. & Achache, J., 1998. Can the thermal thickness of the continental lithosphere be estimated from Magsat data?, *Tectonophysics*, **284**, 19–29.
- Hollerbach, R., 1996. On the theory of the geodynamo, *Phys. Earth planet. Inter.*, **98**, 163–185.
- Holme, R. & Bloxham, J., 1996. The treatment of attitude errors in satellite geomagnetic data, *Phys. Earth planet. Inter.*, **98**, 221–233.
- IAGA Division I Working Group 1, 1981. International Geomagnetic Reference Fields: DGRF 1965, DGRF 1970, DGRF 1975, and IGRF 1980, *EOS, Trans. Am. geophys. Un.*, **57**, 120–121.
- Jackson, A., 1994. Statistical treatment of crustal magnetization, *Geophys. J. Int.*, **119**, 991–998.
- Jackson, A., Jonkers, A.R.T. & Walker, M.R., 2000. Four centuries of geomagnetic secular variation from historical records, *Phil. Trans. R. Soc. Lond., A.*, **358**, 957–990.
- Kaiser, J.F., 1974. Nonrecursive digital filter design using the  $\text{Io-sinh}$  window function, *Proc. IEEE Int. Symp. Circuits Syst.*, **1974**, 20–23.
- Kanasewich, E.R., 1981. *Time sequence analysis in geophysics*, pp. 1–480, The University of Alberta Press, Edmonton.
- Langel, R.A., 1987. The main geomagnetic field, in *Geomagnetism*, Vol. 1, pp. 249–512, ed. Jacobs, J.A., Academic Press, London.
- Langel, R.A., 1990. Global magnetic anomaly maps derived from POGO spacecraft data, *Phys. Earth planet. Inter.*, **62**, 208–230.
- Langel, R.A. & Baldwin, R.T., 1991. Geodynamics branch data base for main magnetic field analysis, *NASA Tech. Mem.*, **104542**, 1–101.
- Langel, R.A. & Estes, R.H., 1982. A geomagnetic field spectrum, *Geophys. Res. Lett.*, **9**, 250–253.
- Langel, R.A. & Estes, R.H., 1985a. Large-scale, near-Earth magnetic fields from external sources and the corresponding induced internal field, *J. geophys. Res.*, **90**, 2487–2494.
- Langel, R.A. & Estes, R.H., 1985b. The near-Earth magnetic field at 1980 determined from Magsat data, *J. geophys. Res.*, **90**, 2495–2510.
- Langel, R.A. & Hinze, W.J., 1998. *The Magnetic Field of the Earth's Lithosphere: The Satellite Perspective*, pp. 1–429, Cambridge University Press, Cambridge.
- Langel, R.A., Coles, R.L. & Mayhew, M.A., 1980. Comparisons of magnetic anomalies of lithospheric origin measured by satellite and airborne magnetometers over western Canada, *Can. J. Earth. Sci.*, **17**, 7.
- Langel, R.A., Estes, R.H. & Mead, G.D., 1982. Some new methods in geomagnetic field modeling applied to the 1960–1980 epoch, *J. Geomag. Geoelectr.*, **34**, 327–349.
- Langel, R.A., Purucker, M.E. & Rajaram, M., 1993. The equatorial electrojet and associated currents as seen in Magsat data, *J. Atmos. Terr. Phys.*, **55**, 1233–1269.
- Langel, R.A., Sabaka, T.J., Baldwin, R.T. & Conrad, J.A., 1996. The near-Earth magnetic field from magnetospheric and quiet-day ionospheric sources and how it is modeled, *Phys. Earth planet. Inter.*, **98**, 235–267.
- Loves, F.J., 1974. Spatial power spectrum of the main geomagnetic field, and extrapolation to the core, *Geophys. J. R. astr. Soc.*, **36**, 717–730.
- Maeda, H., Iyemori, T., Araki, T. & Kamei, T., 1982. New evidence of a meridional current system in the equatorial ionosphere, *Geophys. Res. Lett.*, **9**, 337–340.
- Malin, S.R.C., 1973. Worldwide distribution of geomagnetic tides, *Phil. Trans. Roy. Soc. Lond., A.*, **274**, 551–594.
- Malin, S.R.C. & Gupta, J.C., 1977. The Sq current system during the International Geophysical Year, *Geophys. J. R. astr. Soc.*, **49**, 515–529.
- Malin, S.R.C. & Mete Isikara, A., 1976. Annual variation of the geomagnetic field, *Geophys. J. R. astr. Soc.*, **47**, 445–457.
- Matsushita, S. & Maeda, H., 1965. On the geomagnetic quiet solar daily variation field during the IGY, *J. geophys. Res.*, **70**, 2535–2558.
- Mauersberger, P., 1956. Das Mittel der Energiedichte des geomagnetischen Hauptfeldes an der Erdoberfläche und seine säkulare Änderung, *Gerlands Beitr. Geophys.*, **65**, 207–215.
- Mayaud, P.N., 1972. The  $aa$  indices: a 100-year series characterizing magnetic activity, *J. geophys. Res.*, **77**, 6870.
- Mayaud, P.N., 1980. Derivation, meaning and use of geomagnetic indices, *Geophys. Monogr. Am. geophys. Un.*, **22**.
- Musmann, G. & Seiler, E., 1978. Detection of meridional currents in the equatorial ionosphere, *Geophys. J.*, **44**, 357–372.
- Olsen, N., 1993. The solar cycle variability of lunar and solar daily geomagnetic variations, *Ann. Geophys.*, **11**, 254–262.
- Olsen, N., 1996. Magnetospheric contributions to geomagnetic daily variations, *Ann. Geophys.*, **14**, 538–544.
- Olsen, N., 1997a. Ionospheric  $F$ -region currents at middle and low latitudes estimated from Magsat data, *J. geophys. Res.*, **102**, 4563–4576.
- Olsen, N., 1997b. *Geomagnetic Tides and Related phenomena*, in *Lecture notes in Earth Sciences*, **66**, ed. Wilhelm, H., Springer-Verlag, Berlin.
- Olsen, N., 1998. The electrical conductivity of the mantle beneath Europe derived from C-Responses from 3 h to 720 h, *Geophys. J.*, **133**, 298–308.
- Olsen, N., 1999. Induction studies with satellite data, *Surv. Geophys.*, **20**, 309–340.
- Olsen, N. *et al.* 2000. Ørsted initial field model, *Geophys. Res. Lett.*, **27**, 3607–3610.
- Olson, W.P., 1984. Introduction to the topology of magnetospheric current systems, in *Magnetospheric Currents*, pp. 49–62, ed. Potemra, T.A., *Am. geophys. Un. Monogr.*, **28**.
- Onwumechilli, C.A. & Ozoemena, P.C., 1989. Contours of equatorial electrojet current density, *J. Atmos. Terr. Phys.*, **51**, 163–168.
- Parkinson, W.D. & Hutton, V.R.S., 1989. The electrical conductivity of the Earth, in *Geomagnetism*, Vol. 3, pp. 261–322, ed. Jacobs, J.A., Academic Press, London.

- Peredo, M., Stern, D.P. & Tsyganenko, N.A., 1993. Are existing magnetospheric models excessively stretched?, *J. geophys. Res.*, **98**, 15 343–15 354.
- Potemra, T.A., 1982. Birkeland currents: present understanding and some remaining questions, in *High-latitude Space Plasma Physics*, eds Hultqvist, B. & Hagfors, T., *Nobel Symposium*, **54**, Plenum, New York.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. & Flannery, B.P., 1992. *Numerical recipes in FORTRAN: the Art of Scientific Computing*, 2nd edn, pp. 1–963, Cambridge University Press, Cambridge.
- Purucker, M.E., Sabaka, T.J., Langel, R.A. & Olsen, N., 1997. The missing dimension in Magsat and POGO anomaly studies, *Geophys. Res. Lett.*, **24**, 2909–2912.
- Rangarajan, G.K., 1989. Indices of geomagnetic activity, in *Geomagnetism*, Vol. 3, pp. 323–384, ed. Jacobs, J.A., Academic Press, London.
- Ravat, D.N., Langel, R.A., Purucker, M.E., Arkani-Hamed, J. & Alsdorf, D.E., 1995. Global vector and scalar Magsat magnetic anomaly maps, *J. geophys. Res.*, **100**, 20 111–20 135.
- Richmond, A.D., 1995. Ionospheric electrodynamics using magnetic apex coordinates, *J. Geomag. Geoelectr.*, **47**, 191–212.
- Sabaka, T.J. & Baldwin, R.T., 1993. Modeling the Sq magnetic field from POGO and Magsat satellite and cotemporaneous hourly observatory data: phase I, *Contract Report HSTX/G&G9302*, Hughes STX Corp. for NASA/GSFC Contract NAS5–31 760.
- Sabaka, T.J., Langel, R.A., Baldwin, R.T. & Conrad, J.A., 1997. The geomagnetic field 1900–1995, including the large-scale field from magnetospheric sources, and the NASA candidate models for the 1995 revision of the IGRF, *J. Geomag. Geoelectr.*, **49**, 157–206.
- Sabaka, T.J., Olsen, N. & Langel, R.A., 2000. A comprehensive model of the near-Earth magnetic field: phase 3, NASA/TM-2000-209894, 1–75.
- Schmitz, D.R., Meyer, J. & Cain, J.C., 1989. Modeling the Earth's geomagnetic field to very high degree and order, *Geophys. J.*, **97**, 421–430.
- Schmucker, U., 1985. Magnetic and electric fields due to electromagnetic induction by external sources, in *Landolt-Börnstein, New-Series*, 5/2b, Springer-Verlag, Berlin.
- Schmucker, U., 1999. A spherical harmonic analysis of solar daily variations in the years 1964–65-I. Methods, *Geophys. J. Int.*, **136**, 439–454.
- Schumaker, L.L., 1981. *Spline Functions: Basic Theory*, Wiley, New York.
- Seber, G.A.F. & Wild, C.J., 1989. *Nonlinear Regression*, pp. 1–768, Wiley, New York.
- Takeda, M. & Maeda, H., 1983. F-region dynamo in the evening—interpretation of equatorial  $\Delta D$  anomaly found by Magsat, *J. Atmos. Terr. Phys.*, **45**, 401–408.
- Tarantola, A. & Valette, B., 1982. Generalized non-linear inverse problems solved using the least squares criterion, *Rev. Geophys.*, **20**, 219–232.
- Toutenburg, H., 1982. *Prior Information in Linear Models*, pp. 1–215, Wiley, New York.
- Tsyganenko, N.A., 1987. Global quantitative models of the geomagnetic field in the cislunar magnetosphere for different disturbance levels, *Planet. Space Sci.*, **35**, 1347–1358.
- Tsyganenko, N.A., 1989. A magnetospheric magnetic field model with a warped tail current sheet, *Planet. Space Sci.*, **37**, 5–20.
- Volland, H., 1984. *Atmospheric Electrodynamics*, pp. 1–205, Springer-Verlag, Berlin.
- Walker, A.D. & Backus, G.E., 1997. A six-parameter statistical model of the Earth's magnetic field, *Geophys. J. Int.*, **130**, 693–700.
- Wessel, P. & Smith, W.H.F., 1991. Free software helps map and display data, *EOS, Trans. Am. geophys. Un.*, **72**, 441.
- Winch, D.E., 1981. Spherical harmonic analysis of geomagnetic tides, 1964–1965, *Phil. Trans. Roy. Soc. Lond., A.*, **303**, 1–104.